

REFERENCE: B05

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
10 August 2006 (10.08.2006)

PCT

(10) International Publication Number
WO 2006/084272 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2006/004280

(22) International Filing Date: 6 February 2006 (06.02.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/650,365 4 February 2005 (04.02.2005) US

(71) Applicants (for all designated States except US):
ROSETTA INPHARMATICS LLC [US/US]; 401 Terry
Avenue North, Seattle, Washington 98108 (US), MERCK
& CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway,
New Jersey 07065 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): DAI, Hongyue
[CN/US]; 16814 118th Avenue N.e., Bothell, Washington

98011 (US). FRIEND, Stephen, H [US/US]; 101 W.
Mermaid Lane, Philadelphia, Pennsylvania 19118 (US).
DEUTSCH, Paul [US/US]; 50 Maybury Hill Road,
Princeton, New Jersey 08540 (US).

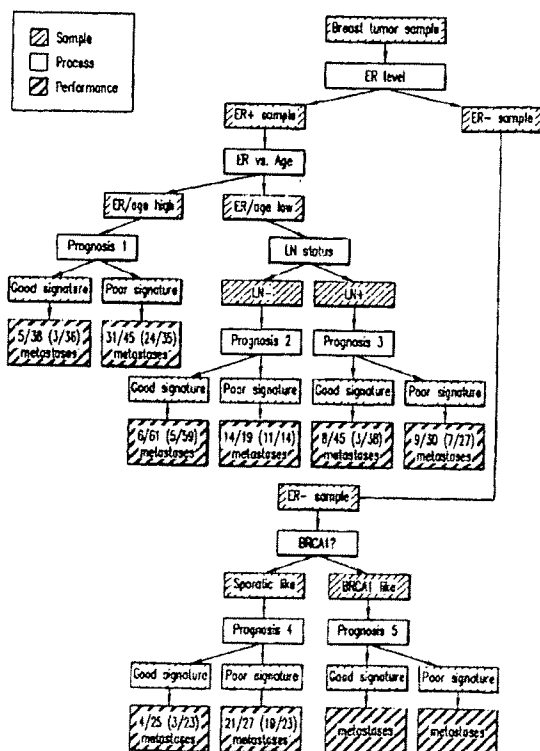
(74) Agents: ANTLEER, Adriane, M. et al.; JONES DAY, 222
East 41st Street, New York, New York 10017-6702 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV,
LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI,
NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG,
SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

[Continued on next page]

(54) Title: METHODS OF PREDICTING CHEMOTHERAPY RESPONSIVENESS IN BREAST CANCER PATIENTS



(57) Abstract: The invention provides a method of predicting the responsiveness of a breast cancer patient to chemotherapy based on the patient's cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from the patient, the patient's estrogen receptor level, and the patient's estrogen receptor level relative to the patient's age. The invention also provides a method for selecting patients for enrollment in a clinical trial of a drug for treating breast cancer based on these factors.

WO 2006/084272 A2



FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

- *without international search report and to be republished upon receipt of that report*

METHODS OF PREDICTING CHEMOTHERAPY RESPONSIVENESS IN BREAST CANCER PATIENTS

This application claims benefit of U.S. Provisional Patent Application No.
60/650,365, filed on February 4, 2005, which is incorporated by reference herein in its
5 entirety.

1. FIELD OF THE INVENTION

The present invention relates to a method of predicting chemotherapy responsiveness
in breast cancer patients. The invention also relates to a method of selecting patients for
10 enrollment in clinical trials of breast cancer drugs.

2. BACKGROUND OF THE INVENTION

The increased number of cancer cases reported in the United States, and, indeed,
around the world, is a major concern. Currently there are only a handful of treatments
available for specific types of cancer, and these provide no guarantee of success. In order to
15 be most effective, these treatments require not only an early detection of the malignancy, but
a reliable assessment of the severity of the malignancy.

The incidence of breast cancer, a leading cause of death in women, has been gradually
increasing in the United States over the last thirty years. Its cumulative risk is relatively high;
1 in 8 women are expected to develop some type of breast cancer by age 85 in the United
20 States. In fact, breast cancer is the most common cancer in women and the second most
common cause of cancer death in the United States. In 1997, it was estimated that 181,000
new cases were reported in the U.S., and that 44,000 people would die of breast cancer
(Parker *et al.*, *CA Cancer J. Clin.* 47:5-27 (1997); Chu *et al.*, *J. Nat. Cancer Inst.* 88:1571-
1579 (1996)). While mechanism of tumorigenesis for most breast carcinomas is largely
25 unknown, there are genetic factors that can predispose some women to developing breast
cancer (Miki *et al.*, *Science*, 266:66-71(1994)).

Sporadic tumors, those not currently associated with a known germline mutation,
constitute the majority of breast cancers. It is also likely that other, non-genetic factors also
have a significant effect on the etiology of the disease. Regardless of the cancer's origin,
30 breast cancer morbidity and mortality increases significantly if it is not detected early in its

progression. Thus, considerable effort has focused on the early detection of cellular transformation and tumor formation in breast tissue.

A marker-based approach to tumor identification and characterization promises improved diagnostic and prognostic reliability. Typically, the diagnosis of breast cancer requires histopathological proof of the presence of the tumor. In addition to diagnosis, histopathological examinations also provide information about prognosis and selection of treatment regimens. Prognosis may also be established based upon clinical parameters such as tumor size, tumor grade, the age of the patient, and lymph node metastasis.

Diagnosis and/or prognosis may be determined to varying degrees of effectiveness by direct examination of the outside of the breast, or through mammography or other X-ray imaging methods (Jatoi, *Am. J. Surg.* 177:518-524 (1999)). The latter approach is not without considerable cost, however. Every time a mammogram is taken, the patient incurs a small risk of having a breast tumor induced by the ionizing properties of the radiation used during the test. In addition, the process is expensive and the subjective interpretations of a technician can lead to imprecision. For example, one study showed major clinical disagreements for about one-third of a set of mammograms that were interpreted individually by a surveyed group of radiologists. Moreover, many women find that undergoing a mammogram is a painful experience. Accordingly, the National Cancer Institute has not recommended mammograms for women under fifty years of age, since this group is not as likely to develop breast cancers as are older women. It is compelling to note, however, that while only about 22% of breast cancers occur in women under fifty, data suggests that breast cancer is more aggressive in pre-menopausal women.

In clinical practice, accurate diagnosis of various subtypes of breast cancer is important because treatment options, prognosis, and the likelihood of therapeutic response all vary broadly depending on the diagnosis. Accurate prognosis, or determination of distant metastasis-free survival could allow the oncologist to tailor the administration of adjuvant chemotherapy, with women having poorer prognoses being given the most aggressive treatment. Furthermore, accurate prediction of poor prognosis would greatly impact clinical trials for new breast cancer therapies, because potential study patients could then be stratified according to prognosis. Trials could then be limited to patients having poor prognosis, in turn making it easier to discern if an experimental therapy is efficacious.

To date, no set of satisfactory predictors for prognosis based on the clinical information alone has been identified. Many have observed that the ER status has a dominant signature in the breast tumor gene expression profiling. See West *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 98:11462 (2001); van 't Veer *et al.*, *Nature* 415:530 (2002); Sorlie *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:8418 (2003); Perou *et al.*, *Nature* 406:747 (2000); Gruvberger *et al.*, *Cancer Res.* 61:5979 (2001); Sotiriou *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:10393 (2003). It is generally accepted that there is some relationship between patient survival and ER status. van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002); Surowiak *et al.*, *Folia Histochem. Cytobiol.* 39:143 (2001); Pichon *et al.*, *Br. J. Cancer* 73:1545 (1996); Collett *et al.*, *J. Clin. Pathol.* 49:920 (1996). *BRCA1* mutations are related to the familial cancer susceptibility. Biesecker *et al.*, *JAMA* 269:1970 (1993); Easton *et al.*, *Cancer Surv.* 18:95 (1993). Age is also considered to be a prognosis factor since young cancer patients tend to have poor tumors. Maggard *et al.*, *J. Surg. Res.* 113:109 (2003). Lymph node status is a factor in deciding the treatment. Eifel *et al.*, *J. Natl. Cancer Inst.* 93:979 (2001).

The discovery and characterization of *BRCA1* and *BRCA2* has recently expanded our knowledge of genetic factors which can contribute to familial breast cancer. Germ-line mutations within these two loci are associated with a 50 to 85% lifetime risk of breast and/or ovarian cancer (Casey, *Curr. Opin. Oncol.* 9:88-93 (1997); Marcus *et al.*, *Cancer* 77:697-709 (1996)). Only about 5% to 10% of breast cancers, however, are associated with breast cancer susceptibility genes, *BRCA1* and *BRCA2*. The cumulative lifetime risk of breast cancer for women who carry the mutant *BRCA1* is predicted to be approximately 92%, while the cumulative lifetime risk for the non-carrier majority is estimated to be approximately 10%. *BRCA1* is a tumor suppressor gene that is involved in DNA repair and cell cycle control, which are both important for the maintenance of genomic stability. More than 90% of all mutations reported so far result in a premature truncation of the protein product with abnormal or abolished function. The histology of breast cancer in *BRCA1* mutation carriers differs from that in sporadic cases, but mutation analysis is the only way to find the carrier. Like *BRCA1*, *BRCA2* is involved in the development of breast cancer, and like *BRCA1* plays a role in DNA repair. However, unlike *BRCA1*, it is not involved in ovarian cancer.

Other genes have been linked to breast cancer, for example c-erb-2 (*HER2*) and p53 (Beenken *et al.*, *Ann. Surg.* 233(5):630-638 (2001). Overexpression of c-erb-2 (*HER2*) and p53 have been correlated with poor prognosis (Rudolph *et al.*, *Hum. Pathol.* 32(3):311-319

(2001), as has been aberrant expression products of *mdm2* (Lukas *et al.*, Cancer Res. 61(7):3212-3219 (2001) and cyclin1 and p27 (Porter & Roberts, International Publication WO98/33450, published August 6, 1998).

The detection of *BRCA1* or *BRCA2* mutations represents a step towards the design of therapies to better control and prevent the appearance of these tumors. Recently, many studies have used gene expression profiling to analyze various cancers, and those studies have provided new diagnosis and prognosis information in the molecular level. See Zajchowski *et al.*, "Identification of Gene Expression Profiled that Predict the Aggressive Behavior of Breast Cancer Cells," *Cancer Res.* 61:5168 (2001); West *et al.*, "Predicting the Clinical Status of Human Breast Cancer by Using Gene Expression Profiles," *Proc. Natl. Acad. Sci. U.S.A.* 98:11462 (2001); van 't Veer *et al.*, "Gene Expression Profiling Predicts the Outcome of Breast Cancer," *Nature* 415:530 (2002); Roberts *et al.*, "Diagnosis and Prognosis of Breast Cancer Patients," WO 02/103320; Sorlie *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:8418 (2003); Perou *et al.*, *Nature* 406:747 (2000); Khan *et al.*, *Cancer Res* 58, 5009 (1998); Golub *et al.*, *Science* 286, 531 (1999); DeRisi *et al.*, *Nat. Genet.* 14:457 (1996); Alizadeh *et al.*, *Nature* 403, 503 (2000). Methods for the identification of informative genesets for various cancers have also been described. See Roberts *et al.*, "Diagnosis and Prognosis of Breast Cancer Patients," WO 02/103320; Golub *et al.*, United States Patent No. 6,647,341.

Genesets have been identified that are informative for differentiating individuals having, or suspected of having, breast cancer based on estrogen receptor (ER) status, or *BRCA1* mutation vs. sporadic (*i.e.*, other than *BRCA1*-type) mutation status. See Roberts *et al.*, WO 02/103320; van't Veer *et al.*, *Nature* 415:530 (2001). Genesets have also been identified that enable the classification of sporadic tumor-type individuals as those who will likely have no metastases within five years of initial diagnosis (*i.e.*, individuals with a good prognosis) or those who will likely have a metastasis within five years of initial diagnosis (*i.e.*, those having a poor prognosis). Roberts, *supra*; van't Veer, *supra*.

Roberts *et al.* WO 02/103320 and van de Vijver *et al.*, *N. Engl. J. Med.* 347:1999 (2002) describe a 70-gene set, useful for the prognosis of breast cancer, which outperformed clinical measures of prognosis, and which showed good potential in selecting good outcome patients, thereby avoiding over-treatment. The expression of genes with most predictive

value, however, was not homogeneous among poor patients, suggesting the need for improvement.

Although the patterns of gene expression as described in Roberts *et al.* were correlated with existing clinical indicators such as estrogen receptor and *BRCA1* status, clinical measures were not incorporated. Furthermore, although the poor-outcome group in particular showed heterogeneity in expression pattern, the best classifier decision rule found during these studies was a fairly simple one based on the similarity of a patient profile to the average profile of a good-outcome training group.

It is evident that breast cancer is the result of more than one type of molecular event.

Likewise, a variety of other conditions, such as other cancers; non-cancer diseases such as diabetes, autoimmune or neurodegenerative disorders, obesity; etc., are also the result of more than one molecular event. Moreover, an individual's response to exposure to particular environmental conditions, for example, exposure to natural or man-made agents, such as toxins, pollutants, drugs, food additives, etc., likely result from more than one molecular event. Thus, there exists a need for improved prognostic methods so that appropriate courses of prophylaxis and/or therapy may be provided. Genesets having improved prognostic power can be identified by first identifying discrete subsets of individuals based on genotypic or phenotypic characteristics relevant to the disease or condition, and then identifying genesets informative for prognosis within those subsets of patients. Individuals having the condition, or who are suspected of having the condition, such as breast cancer, would then be provided therapies appropriate to the molecular mechanisms underlying the condition. The present invention provides such methods for breast cancer, and for other cancers, diseases or conditions.

3. SUMMARY OF THE INVENTION

The present invention provides a method for predicting responsiveness of a breast cancer patient to chemotherapy based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method

comprising predicting said patient to exhibit (a) enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is low; or (b) reduced response to chemotherapy as compared to patients in the general
5 population of breast cancer patients if (i) said ER level is ER⁻, or (ii) said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is high; wherein said cellular constituent profile is classified as a poor prognosis profile if said cellular constituent profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template, said good prognosis template comprising
10 measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of good outcome patients and said poor prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of poor outcome patients, wherein a good outcome patient is a breast cancer patient who has non-occurrence of metastases within a first period
15 of time after initial diagnosis and a poor outcome patient is a patient who has occurrence of metastases within a second period of time after initial diagnosis.

In one embodiment, said cellular constituent profile is determined as a poor prognosis profile, and said ER level is determined as ER⁺, and said ER/AGE is determined as low, and said patient is predicted to have enhanced response to chemotherapy as compared to patients
20 in the general population of breast cancer patients. In another embodiment, said cellular constituent profile is determined as a poor prognosis profile, said ER level is determined as ER⁺, and said ER/AGE is determined as high, and said patient is predicted to have reduced response to chemotherapy as compared to patients in the general population of breast cancer patients. In still another embodiment, said ER level is determined as ER⁻, and said patient is
25 predicted to have reduced response to chemotherapy as compared to patients in the general population of breast cancer patients.

The invention also provides a method for selecting a patient for enrollment in a clinical trial of a drug for treating breast cancer based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in
30 a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said

ER/AGR being a metric of said ER level relative to the age of said patient, said method comprising selecting a patient for inclusion in said clinical trial if (a) said ER level is ER⁻; or (b) said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high; wherein said cellular constituent profile is classified as a poor prognosis
5 profile if said cellular constituent profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template, said good prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of good outcome patients and said poor prognosis template
10 comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of poor outcome patients, wherein a good outcome patient is a breast cancer patient who has non-occurrence of metastases within a first period of time after initial diagnosis and a poor outcome patient is a breast cancer patient who has occurrence of metastases within a second period of time after initial diagnosis.

In one embodiment, said ER level is determined as ER⁻, and said patient is selected.
15 In another embodiment, said cellular constituent profile is determined as a poor prognosis profile, said ER level is determined as ER⁺, and said ER/AGE is determined as high, and said patient is selected.

The invention also provides a method for identifying a breast cancer patient as a good or poor candidate for chemotherapy based on one or more of the following (i) a cellular
20 constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method
25 comprising (a) determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, said breast cancer patient is identified as a good candidate for chemotherapy; or (b) determining whether said ER level is ER⁻, whereby if said ER level is ER⁻, said the breast cancer patient is
30 identified as a poor candidate for chemotherapy; or (c) determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺,

and said ER/AGE is high, said breast cancer patient is identified as a poor candidate for chemotherapy.

In one embodiment, said cellular constituent profile is determined as a poor prognosis profile, said ER level is determined as ER⁺, and said ER/AGE is determined as low, and said
5 breast cancer patient is identified as a good candidate for chemotherapy. In another embodiment, said ER level is determined as ER⁻, and said the breast cancer patient is identified as a poor candidate for chemotherapy. In still another embodiment, said cellular constituent profile is determined as a poor prognosis profile, said ER level is determined as ER⁺, and said ER/AGE is determined as high, and said breast cancer patient is identified as a
10 poor candidate for chemotherapy.

In embodiments of the methods of the invention, said first period is 10 years and said second period is 10 years. In another embodiment, said patient is under the age of 55, and wherein each said good outcome patient and each said poor outcome patient are under the age of 55 at time of diagnosis of breast cancer.

15 In one embodiment, any one of the methods of the invention further comprises determining said cellular constituent profile, said ER level, and/or, said ER/AGE.

In another embodiment, in any one of the methods of the invention said cellular constituent profile comprises measurements of a plurality of transcripts in a sample derived from said patient, wherein said good prognosis template comprises measurements of said
20 plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

In one embodiment, said cellular constituent profile is a differential expression profile
25 comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample. In one embodiment, said differential measurements is selected from the group consisting of xdev, log(ratio), error-weighted log(ratio), and mean subtracted log(intensity).

In another embodiment, said cellular constituent profile comprises measurements of a
30 plurality of protein species in a sample derived from said patient, wherein said good

prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of poor outcome patients.

5 In one embodiment, measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of good outcome patients.

 In one embodiment, similarity of said cellular constituent profile to said good prognosis template is represented by a correlation coefficient between said cellular
10 constituent profile and said good prognosis template, wherein said correlation coefficient greater than a correlation threshold, e.g., 0.5, indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

 In another embodiment, similarity of said cellular constituent profile to said good prognosis template is represented by a distance between said cellular constituent profile and
15 said good prognosis template, wherein said distance less than a given value indicates a high similarity and said distance equal to or greater than said given value indicates a low similarity.

 In one embodiment, said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor in said patient relative to expression level of said gene
20 in said control sample, and wherein said ER level is classified as ER⁺ if $\log_{10}(\text{ratio})$ of said expression level is greater than -0.65, and wherein said ER level is classified as ER⁻ if $\log_{10}(\text{ratio})$ of said expression level is equal to or less than -0.65. In one embodiment, said gene encoding said estrogen receptor is the estrogen receptor α gene.

 In one embodiment, said ER/AGE is classified as high if said ER level as measured
25 by $\log_{10}(\text{ratio})$ is greater than $c \cdot (\text{AGE} - d)$, and wherein said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

 In one embodiment, said estrogen receptor level is measured by an oligonucleotide probe that detects a transcript corresponding to the gene having accession number

NM_000125, wherein said control sample is a pool of breast cancer cells of different patients, and wherein $c = 0.1$ and $d = 42.5$.

In one embodiment, said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients. In another embodiment,
5 said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript of said gene encoding said estrogen receptor.

In one embodiment, said plurality of transcripts consists of transcripts corresponding to at least a portion of the set of genes listed in Table 8. In another embodiment, said plurality of transcripts consists of all transcripts corresponding to genes listed in Table 8.

10 In one embodiment, said chemotherapy is carried out using the CMF combination consisting of cyclophosphamide, methotrexate, and 5-fluorouracil.

The invention also provides a computer system comprising a processor, and a memory coupled to said processor and encoding one or more programs that cause the processor to carry out any one of the methods of the invention.

15 The invention also provides computer program product for use in conjunction with a computer having a processor and a memory connected to the processor, said computer program product comprising a computer readable storage medium having a computer program mechanism encoded thereon, which computer program mechanism may be loaded
20 into the memory of said computer and cause said computer to carry out any one of the methods of the invention.

The invention also provides a method for predicting responsiveness of a breast cancer patient to chemotherapy based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor
25 transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method comprising predicting said patient to exhibit (a) enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular
30 constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is

low; or (b) reduced response to chemotherapy as compared to patients in the general population of breast cancer patients if (i) said ER level is ER⁻, or (ii) said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is high, wherein said cellular constituent profile is classified as a good prognosis profile if said
5 cellular constituent profile predicts non-occurrence of metastases in said breast cancer patient within a predetermined period of time after initial diagnosis, and wherein said cellular constituent profile is classified as a poor prognosis profile if said cellular constituent profile predicts occurrence of metastasis within said predetermined period of time. In one embodiment, the predetermined period of time is 5 or 10 years.

10 In one embodiment, said cellular constituent profile comprises measurements of a plurality of transcripts in a sample derived from said patient. In another embodiment, said cellular constituent profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus
15 measurements of said plurality of transcripts in a control sample. In one embodiment, said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts.

In a preferred embodiment, said plurality of transcripts consists of transcripts corresponding to at least a portion, e.g., at least 5, 10, 40, 50 or all, of the genes listed in Table 8.

20 In one embodiment, said cellular constituent profile is classified as a good prognosis profile or a poor prognosis profile using an artificial neural network, which receives said cellular constituent profile as an input and generates an output comprising data indicating whether said cellular constituent profile is a good prognosis profile or a poor prognosis profile.

25 In another embodiment, said cellular constituent profile is classified as a good prognosis profile or a poor prognosis profile using a supporting vector machine, which receives said cellular constituent profile as an input and generates an output comprising data indicating whether said cellular constituent profile is a good prognosis profile or a poor prognosis profile.

30 4. BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 depicts the decision tree that resulted in the five patient subsets used to identify informative prognosis-related genes.

FIG. 2 shows the relationship between ER level and age. (A) Scatter plot of ER vs. age for ER+ patients. Black dots indicate metastases free samples, and gray dots indicate metastases samples. It appears that patients of ER+ group can be subdivided into “ER+, ER/AGE high” group (above the black line) and “ER+, ER/AGE low” (below the black line) group. The black line is approximated by $ER = 0.1 * (AGE - 42.5)$, and the dashed line by $ER = 0.1 * (age - 50)$. Within each population, the ER level also increases with age. (B) Age distribution of all patients in ER+ samples. A bimodal distribution is observed. (C) ER-modulated age ($age - 10*$) distribution of all patients in ER+ samples. A bimodal distribution is observed. (D) Age distribution of samples with metastasis. (E) ER-modulated age distribution of samples with metastasis. The three peaks appearing in this distribution suggest a polymorphism.

FIG. 3. Performance of classifier for the “ER⁻/sporadic” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

FIG. 4. Performance of classifier for the “ER+, ER/AGE high” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, and squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

FIG. 5. Performance of classifier for the “ER+, ER/AGE low/LN⁻” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot

between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases-free samples, and squares indicate samples with metastases. Dashed line indicates the threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

FIG. 6. Performance of classifier for the “ER+, ER/AGE low/LN+” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

FIG. 7. Performance of classifier for the “ER⁻, BRCA1” group. (A) Error rate obtained from leave-one-out cross validation (LOOCV) for predicting the disease outcome as a function of the number of reporter genes used in the classifier. (B) Scatter plot between correlation to good group (X axis) and to poor group (Y axis). Circles indicate metastases free samples, squares indicate samples with metastases. Dashed line: threshold for separating poor from good. (C) Error rate calculated with respect to good outcome group (good outcome misclassified as poor divided by total number of good), or poor outcome group (poor outcome misclassified as good divided by total number of poor), or the average of the two rates.

FIG. 8. Heatmaps of genes representing key biological functions in subgroups of patients: A: Cell cycle genes are predictive of outcome in patients with ER/age high. B: Cell cycle genes are not predictive of outcome in “ER- and sporadic” patients C: Glycolysis genes are predictive of outcome in patients with ER/age low and LN-. D: Glycolysis genes are not predictive of outcome in “ER- & BRCA1” patients.

FIG. 9 shows metastasis free probability for the entire set of breast cancer patients younger than 55 years of age, with or without chemotherapy.

FIG. 10 shows the effect of chemotherapy in (a) patients with “good signature” as predicted by the “70 gene” classifier, and (b) patients with “poor signature”.

FIG. 11 shows the effect of chemotherapy in (a) ER- patients, and (b) ER+ patients with “poor signature”.

5 FIG. 12 ER+ patients are divided into “ER/age high” (above the line) and “ER/age low” (below the line) group according to their ER expression level relative to patient age at diagnosis. The ER expression level is measured by gene expression array. The control channel is the pool of all breast cancer samples. The line is represented as $y = 0.1 * (x - 42.5)$.

10 FIG. 13 shows the effect of chemotherapy in “poor signature” patients in (a) “ER/age high” group, and (b) “ER/age low” group.

5. DETAILED DESCRIPTION OF THE INVENTION

The invention provides a method of predicting the responsiveness of a breast cancer patient to chemotherapy based on the patient’s cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from the patient, the patient’s estrogen receptor level (hereinafter designated as “ER level”), and the patient’s estrogen receptor level relative to the patient’s age (hereinafter designated as “ER/AGE”). In the method of the invention, a patient is classified into an appropriate chemotherapy responsiveness group as compared to patients in the general population of breast cancer patients. The inventors have discovered that, among patients whose cellular constituent profile indicates poor prognosis, a patient’s responsiveness to chemotherapy depends not only on the patient’s ER level, but also on the change of the ER level with age. The inventors have found that patients who show high ER level at an earlier age (thus a high ER/AGE) show little response to chemotherapy, whereas patients who show high ER level at later age (thus a low ER/AGE) show increased response to chemotherapy. In particular, the inventors have found that a patient exhibits enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if the patient’s cellular constituent profile indicates poor prognosis, and the patient’s ER level is ER⁺, and the patient’s estrogen receptor level relative to the patient’s age ER/AGE is low, whereas a patient exhibits reduced response to chemotherapy as compared to patients in the general population if (i) the patient’s ER level is ER⁻, or (ii) the patient’s cellular constituent profile indicates poor prognosis, and the patient’s ER level is ER⁺, and the patient’s estrogen receptor level relative to the patient’s

age ER/AGE is high. Herein, ER⁺ designates a high ER level and ER⁻ designates a low ER level. In addition, the inventors have also found that if a patient's cellular constituent profile indicates a good prognosis, then the patient does not need chemotherapy. The method of the invention is particularly useful for predicting the responsiveness of a breast cancer patient
5 under the age of 55.

The invention also provides a method for selecting patients for enrollment in a clinical trial of a drug for treating breast cancer based on each patient's cellular constituent profile, the patient's estrogen receptor level, and the patient's estrogen receptor level relative to the patient's age. In the method of the invention, patients having reduced responsiveness to
10 chemotherapy as compared to patients in the general population of breast cancer patients are preferably selected for inclusion in a clinical trial. In one embodiment, a patient who has either ER⁻ or a combination of (a) poor prognosis, (b) ER⁺, and (c) ER/AGE high are selected for inclusion in the trial. Such patients are predicted not to exhibit significant response to chemotherapy, and thus are good candidates in determining efficacy of a new breast cancer
15 drug.

Patient groups can be classified according to at least one of age, lymph node (LN) status, estrogen receptor (ER) level, and *BRCA1* mutation status into discrete patient subsets. These clinical factors have been implicated in tumor etiology as well as differences in disease outcome. These characteristics are not limiting; other genotypic or phenotypic characteristics
20 of breast cancer, for example, tumor grade, tumor size, tumor cell type, etc., may be used, alone or in combination with those listed herein, in order to classify individuals. The differences in gene expression or in tumor fate related to these parameters likely represent differences in tumor origin and tumor genesis, and are therefore good candidates for tumor stratification. Genesets informative for prognosis within each subset are then identified.
25 New breast cancer patients are then classified using the same criteria, and a prognosis is made based on the geneset specific for the patient subset into which the patient falls. In the process of constructing a prognosis classifier within each patient subset, particular attention is paid to the homogeneous patterns related to the tumor outcome. Emergence of such homogeneous prognosis patterns may indicate the most common mechanism to metastasis within a subset.
30 At the same time, successful identification of such patterns also justifies the parameters being used for the tumor stratification. To differentiate this approach from an mRNA-alone

approach, we refer to the current approach of integrating clinical data with the gene expression data as a “comprehensive prognosis”.

The measurements in the cellular constituent profiles can be any suitable measured values of the cellular constituents, e.g., measurements of the expression levels of genes. The measurement of the expression level of a gene can be direct or indirect, e.g., directly of abundance levels of RNAs or proteins or indirectly, by measuring abundance levels of cDNAs, amplified RNAs or DNAs, proteins, or activity levels of RNAs or proteins, or other molecules (e.g., a metabolite) that are indicative of the foregoing. In one embodiment, the profile comprises measurements of abundances of the transcripts of the marker genes. The measurements of transcripts can be direct (of the transcripts themselves) or indirect (of, e.g., their cDNAs). The measurement of abundance can be a measurement of the absolute abundance of a gene product. The measurement of abundance can also be a value representative of the absolute abundance, e.g., a normalized abundance value (e.g., an abundance normalized against the abundance of a reference gene product) or an averaged abundance value (e.g., average of abundances obtained at different time points or from different tumor cell samples from the patients, or average of abundances obtained using different probes, etc.), or a combination of both. As an example, the measurement of abundance of a gene transcript can be a value obtained using an Affymetrix® GeneChip® to measure hybridization to the transcript.

In another embodiment, the expression profile is a differential expression profile comprising differential measurements of a plurality of transcripts in a sample derived from the patient versus measurements of the plurality of transcripts in a reference sample, e.g., a cell sample of normal cells. Each differential measurement in the profile can be but is not limited to an arithmetic difference, a ratio, or a log(ratio). As an example, the measurement of abundance of a gene transcript can be a value for the transcript obtained using a cDNA or ink-jet polynucleotide array in a two-color measurement.

As used herein, “*BRCA1* tumor” or “*BRCA1* type” means a tumor having cells containing a mutation of the *BRCA1* locus.

A “patient subset” is a group of individuals, all of whom have a particular condition, or are subject to a particular condition, which is distinguished from other individuals having that condition by one or more phenotypic, genotypic or clinical characteristics of the

condition, or of a response to the condition. For example, where the condition is breast cancer, individuals may belong to an "ER⁺" or an "ER⁻" patient subset, or may belong to a particular age group patient subset.

5 A gene and/or marker is "informative" for a condition, phenotype, genotype or clinical characteristic if the expression of the gene or marker is correlated or anticorrelated with the condition, phenotype, genotype or clinical characteristic to a greater degree than would be expected by chance.

10 The patient's estrogen receptor level relative to the patient's age can be measured using an appropriate metric designated as ER/AGE. An individual of a given age can be classified as "ER/AGE high" if the individual's ER level is higher than a threshold value for the given age. The threshold can be age-dependent, i.e., a different threshold for each different age. In one embodiment, the age-dependent threshold value is calculated as $c \cdot (AGE - d)$, where c is a coefficient, AGE is the age of the patient, and d is an age threshold. The parameters c and d depend on the particular measures and/or units of the ER level and
15 AGE. They can be determined by fitting patients' ER level-age distribution to a bimodal distribution of two subgroups each having a different ER level-age dependence. In a specific embodiment, $c = 0.1$ and $d = 42.5$ is used for ER levels represented by a log(ratio) of ER expression level. Thus, for example, the threshold for a 45-year old individual in this embodiment is 0.1 (45-42.5), or 0.25, and if the log(ratio) of ER expression level of the
20 individual is equal to or greater than 0.25, the individual is classified as "ER/AGE high"; otherwise, the individual is classified as "ER/AGE low."

5.1 IDENTIFICATION OF DIAGNOSTIC AND PROGNOSTIC MARKER SETS

The sets of genes and/or markers that can be used in conjunction with the present invention for diagnosis and/or prognosis of breast cancer can be identified using methods
25 described in this section. In preferred embodiments, the method involves first stratifying breast cancer patients according to phenotypic, genotypic and/or clinical parameters into subsets, and then identifying markers that discriminate diagnosis and/or prognosis in each subsets.

5.1.1 IDENTIFICATION OF CONDITION SUBSETS

30 In one embodiment, the subsets are distinguished by phenotypic, genotypic, and/or clinical characteristics of breast cancer. In this embodiment, groups of individuals are

classified according to one or more phenotypic, genotypic, or clinical characteristics of breast cancer into different patient subsets. At any step in the process of subdividing a patient population into patient subsets, the expression level of one or more genes may be determined in order to identify whether a prognosis-informative set of genes may be identified for the

5 particular patient subset. If a gene set is identified for a subset of patients, but is not as informative as desired, the patient subset may be further divided and a new geneset identified. These subsets may be further subdivided. For example, a group of individuals affected by breast cancer may be classified first on the basis of a phenotypic, genotypic or clinical characteristic A into subsets S1 and S2. The levels of expression of a plurality of genes are

10 then determined in tumor samples taken from individuals that fall within subsets S1 or S2 in order to identify sets of genes informative for prognosis within these subsets. Subsets S1 and S2 may then each be subdivided into two or more subsets based on other phenotypic, genotypic or clinical characteristics. The basis for subdivision, if performed, need not be the same for S1 and S2. For example, in various embodiments, S1 is not subdivided, while S2 is

15 subdivided on the basis of characteristic B; or S1 is subdivided based on characteristic B while S2 is not subdivided; or S1 and S2 are both subdivided on the basis of characteristic B; or S1 is subdivided based on characteristic B, while S2 is subdivided according to characteristic C; and so on. For a particular decision matrix leading to a plurality of patient subsets, the preferred outcome is a prognosis-informative set of genes for each patient subset.

20 Different decision matrices may lead to different patient subsets, which, in turn, may result in different sets of prognosis-informative genes.

In the specific example, a plurality of phenotypic, genotypic or clinical indications are used to classify a breast cancer patient as being a member of one of a plurality of patient subsets, wherein the subsets are medically, biochemically or genetically relevant to breast

25 cancer. For example, a group of patients may be classified into patient subsets based on criteria including, but not limited to, estrogen receptor (ER) status, type of tumor (*i.e.*, *BRCA1*-type or sporadic), lymph node status, grade of cancer, invasiveness of the tumor, or age. "BRCA1-type" indicates that the *BRCA1* mutation is present. In each classification step, a group of cancer patients may be classified into only two classes, for example, ER+ or

30 ER⁻, or into three or more subsets (for example, by tumor grade), depending upon the characteristic used to determine the subsets. As used herein, "ER+" indicates that the estrogen receptor is expressed at some level; for example, it may indicate that the estrogen receptor is detectably expressed, or may indicate that more than 10% of cells may be

histologically stained for the receptor, etc. Conversely, “ER–” indicates that the estrogen receptor is expressed at a reduced level or not at all; for example, it may indicate that the receptor is not detectably expressed, or that 10% or less of cells may be histologically stained for the receptor, etc. Marker gene sets optimized for each phenotypic class are preferably
5 determined after the subsets are established. Where informative markers for a particular patient subset, distinguished from another subset by a particular characteristic of the condition of interest, cannot be determined, the subset may be further divided by another characteristic of the condition to create a plurality of second patient subsets, whereupon genes informative for these second patient subsets may be identified.

10 FIG. 1 depicts the process, described in the Examples, of subdivision of a collection of breast cancer patients according to phenotypic and genotypic characteristics relevant to breast cancer, in preparation for identification of genes informative for prognosis. A collection of breast cancer tumor samples was first subdivided by estrogen receptor status. ER status was chosen because the presence or absence of the estrogen receptor greatly
15 influences the expression of other genes. In the ER+ patient subset, it was noted that patients appeared to be bimodally distributed by ER level vs. age; that is, ER level dependence upon age tended to fall within two classes, as separated by the solid line in FIG. 2A. This bimodality was used to further subdivide ER+ individuals into “ER+, ER/AGE high” individuals and “ER+, ER/AGE low” individuals. A set of informative genes was identified
20 for the ER+, ER/AGE high patient subset. An informative set was not identified for the ER+, ER/AGE low subset, however, so the subset of patients was further divided into LN+ and LN– individuals.

5.1.2 IDENTIFICATION OF MARKER SETS INFORMATIVE FOR PATIENT SUBSETS

25 .Once a patient subset is identified, markers, such as genes, informative for a particular outcome, such as prognosis, may be identified. The methods make use of measured cellular constituent profiles, e.g., expression profiles of a plurality of genes (e.g., measurements of abundance levels of the corresponding gene products), in tumor samples from a plurality of patients in the patient subset whose prognosis outcomes are known. The prognosis outcomes
30 can be the prognosis at a predetermined time after initial diagnosis. The predetermined time can be any convenient time period, e.g., 2, 3, 4, or 5 years. Prognosis markers can be obtained by identifying genes whose expression levels correlate with prognosis outcome, e.g.,

genes whose expression levels in good prognosis patients group are significantly different from those in poor prognosis patients. In preferred embodiments, the tumor samples from the plurality of patients are separated into a good prognosis group and a poor prognosis group for the predetermined time period. Genes whose expression levels exhibit differences between
5 the good and poor prognosis groups to at least a predetermined level are selected as the genes whose expression levels correlate with patient prognosis. This section describes embodiments which employ genes and gene-derived nucleic acids as markers. However, it will be understood by a person skilled in the art that proteins or other cellular constituents may also be used as markers.

10 In a preferred embodiment, the expression profile is a differential expression profile. Each measurement in the profile is a differential expression level of a marker in a breast tumor sample versus that in a reference sample (also termed a standard or control sample). In one embodiment, the reference sample comprises polynucleotide molecules, derived from one or more samples from a plurality of normal individuals. For example, the normal
15 individuals may be persons not afflicted with breast cancer. The standard or control may alternatively comprise polynucleotide molecules, derived from one or more samples derived from individuals having a different form or stage of breast cancer; a different disease or different condition, or individuals exposed or subjected to a different condition, than the individual from which the sample of interest was obtained. The reference or control may be a
20 sample, or set of samples, taken from the individual at an earlier time, for example, to assess the progression of a condition, or the response to a course of therapy.

In a preferred embodiment, the standard or control is a pool of target polynucleotide molecules derived from a plurality of different individuals. However, where protein levels, or the levels of any other relevant biomolecule, are to be compared, the pool may be a pool of
25 proteins or the relevant biomolecule. In a preferred embodiment in the context of breast cancer, the pool comprises samples taken from a number of individuals having sporadic-type tumors.

In another preferred embodiment, the pool comprises an artificially-generated population of nucleic acids designed to approximate the level of nucleic acid derived from
30 each marker found in a pool of marker-derived nucleic acids derived from tumor samples. In another embodiment, the pool, also called a "mathematical sample pool," is represented by a set of expression values, rather than a set of physical polynucleotides; the level of expression

of relevant markers in a sample from an individual with a condition, such as a disease, is compared to values representing control levels of expression for the same markers in the mathematical sample pool. Such a control may be a set of values stored on a computer. Such artificial or mathematical controls may be constructed for any condition of interest.

5 In another embodiment, the reference sample is derived from a normal breast cell line or a breast cancer cell line. In a preferred embodiment, the reference sample comprises samples taken from individuals within a specific patient subset, *e.g.*, "ER+, ER/AGE high" individuals, wherein each of said individuals has a good prognosis, or each of said individuals has a poor prognosis. Of course, where, for example, expressed proteins are used as markers,
10 the proteins are obtained from the individual's sample, and the standard or control can be a pool of proteins from a number of normal individuals, or from a number of individuals having a particular state of a condition, such as a pool of samples from individuals having a particular prognosis of breast cancer.

 The comparison may be accomplished by any means known in the art. For example,
15 expression levels of various markers may be assessed by separation of target polynucleotide molecules (*e.g.*, RNA or cDNA) derived from the markers in agarose or polyacrylamide gels, followed by hybridization with marker-specific oligonucleotide probes. Alternatively, the comparison may be accomplished by the labeling of target polynucleotide molecules followed by separation on a sequencing gel. Polynucleotide samples are placed on the gel
20 such that patient and control or standard polynucleotides are in adjacent lanes. Comparison of expression levels is accomplished visually or by means of densitometer. In a preferred embodiment, the expression of all markers is assessed simultaneously by hybridization to a microarray. In each approach, markers meeting certain criteria are identified as informative for the prognosis of breast cancer.

25 In one embodiment, genes are first screened based on significant variation in expression in a set of breast cancer tumor samples as compared to a standard or control sample. Genes may be screened, for example, by determining whether they show significant variation in at least some samples among the set of samples. Genes that do not show significant variation in at least some samples in the set of samples are presumed not to be
30 informative, and are discarded from further consideration. Genes showing significant variation in at least some samples in the sample set are retained as candidate informative genes. The degree of variation in expression of a gene may be estimated by determining a

difference or ratio of the expression of the gene in a sample and a control. The difference or ratio of expression may be further transformed, *e.g.*, by a linear or log transformation.

Selection of candidate markers may be made based upon either significant up- or down-regulation of the gene in at least some samples in the set or based on the statistical

5 significance (*e.g.*, the *p*-value) of the variation in expression of the gene. Preferably, both selection criteria are used. Thus, in one embodiment of the present invention, genes showing both a more than two-fold change (increase or decrease) in expression as compared to a standard in at least three samples, and a *p*-value of variation in expression of the gene in the set of tumor samples as compared to the standard sample is no more than 0.01 (*i.e.*, is
10 statistically significant) are selected as candidate genes associated with prognosis of breast cancer in a patient subset.

In the present invention, a “good prognosis” predicts non-occurrence of metastases within a predetermined period of time, *e.g.*, 1, 2, 3, 4, 5 or more years, after initial diagnosis, whereas a “poor prognosis” predicts occurrence of metastasis within that period. In a specific
15 embodiment, a “good prognosis” predicts non-occurrence of metastases within 5 years after initial diagnosis, and a “poor prognosis” predicts occurrence of metastasis within that period.

Expression profiles comprising a plurality of different genes in a plurality of *N* breast cancer tumor samples can be used to identify markers that correlate with, and therefore are useful for discriminating, different clinical categories. In a specific embodiment, a
20 correlation coefficient ρ between a vector \vec{c} representing clinical categories or clinical parameters, *e.g.*, a good or poor prognosis, in the *N* tumor samples and a vector \vec{r} representing the measured expression levels of a gene in the *N* tumor samples is used as a measure of the correlation between the expression level of the gene and clinical category. The expression levels can be a measured abundance level of a transcript of the gene, or any
25 transformation of the measured abundance, *e.g.*, a logarithmic or a log ratio. Specifically, the correlation coefficient may be calculated as:

$$\rho = (\vec{c} \cdot \vec{r}) / (\|\vec{c}\| \cdot \|\vec{r}\|) \quad (1)$$

Markers for which the coefficient of correlation exceeds a cutoff are identified as prognosis-informative markers specific for a particular clinical category, *e.g.*, a good
30 prognosis, within a given patient subset. Such a cutoff or threshold may correspond to a certain significance of the set of obtained discriminating genes. The threshold may also be

selected based on the number of samples used. For example, a threshold can be calculated as $3 \times 1/\sqrt{n-3}$, where $1/\sqrt{n-3}$ is the distribution width and n = the number of samples. In a specific embodiment, markers are chosen if the correlation coefficient is greater than about 0.3 or less than about -0.3.

5 Next, the significance of the set of marker genes can be evaluated. The significance may be calculated by any appropriate statistical method. In a specific example, a Monte-Carlo technique is used to randomize the association between the expression profiles of the plurality of patients and the clinical categories to generate a set of randomized data. The same marker selection procedure as used to select the marker set is applied to the randomized
10 data to obtain a control marker set. A plurality of such runs can be performed to generate a probability distribution of the number of genes in control marker sets. In a preferred embodiment, 10,000 such runs are performed. From the probability distribution, the probability of finding a marker set consisting of a given number of markers when no correlation between the expression levels and phenotype is expected (i.e., based randomized
15 data) can be determined. The significance of the marker set obtained from the real data can be evaluated based on the number of markers in the marker set by comparing to the probability of obtaining a control marker set consisting of the same number of markers using the randomized data. In one embodiment, if the probability of obtaining a control marker set consisting of the same number of markers using the randomized data is below a given
20 probability threshold, the marker set is said to be significant.

Once a marker set is identified, the markers may be rank-ordered in order of correlation or significance of discrimination. One means of rank ordering is by the amplitude of correlation between the change in gene expression of the marker and the specific condition being discriminated. Another, preferred, means is to use a statistical metric. In a specific
25 embodiment, the metric is a t-test-like statistic:

$$t = \frac{(\langle x_1 \rangle - \langle x_2 \rangle)}{\sqrt{[\sigma_1^2(n_1 - 1) + \sigma_2^2(n_2 - 1)] / (n_1 + n_2 - 1) / (1/n_1 + 1/n_2)}} \quad (2)$$

In this equation, $\langle x_1 \rangle$ is the error-weighted average of the log ratio of transcript expression measurements within a first clinical group (e.g., good prognosis), $\langle x_2 \rangle$ is the error-weighted average of log ratio within a second, related clinical group (e.g., poor prognosis), σ_1 is the

variance of the log ratio within the first clinical group (*e.g.*, good prognosis), n_1 is the number of samples for which valid measurements of log ratios are available, σ_2 is the variance of log ratio within the second clinical group (*e.g.*, poor prognosis), and n_2 is the number of samples for which valid measurements of log ratios are available. The t -value represents the variance-compensated difference between two means. The rank-ordered marker set may be used to optimize the number of markers in the set used for discrimination.

A set of genes for prognosis of breast cancer can also be identified using an iterative approach. This is accomplished generally in a "leave one out" method as follows. In a first run, a subset, for example five, of the markers from the top of the ranked list is used to generate a template, where out of N samples, $N-1$ are used to generate the template, and the status of the remaining sample is predicted. This process is repeated for every sample until every one of the N samples is predicted once. In a second run, one or more additional markers, for example five additional markers, are added, so that a template is now generated from 10 markers, and the outcome of the remaining sample is predicted. This process is repeated until the entire set of markers is used to generate the template. For each of the runs, type 1 error (false negative) and type 2 errors (false positive) are counted. The set of top-ranked markers that corresponds to lowest type 1 error rate, or type 2 error rate, or preferably the total of type 1 and type 2 error rate is selected.

For prognostic markers, validation of the marker set may be accomplished by an additional statistic, a survival model. This statistic generates the probability of tumor distant metastases as a function of time since initial diagnosis. A number of models may be used, including Weibull, normal, log-normal, log logistic, log-exponential, or log-Rayleigh (Chapter 12 "Life Testing", S-PLUS 2000 GUIDE TO STATISTICS, Vol. 2, p. 368 (2000)). For the "normal" model, the probability of distant metastases P at time t is calculated as

$$P = \alpha \times \exp\left(-t^2/\tau^2\right) \quad (3)$$

where α is fixed and equal to 1, and τ is a parameter to be fitted and measures the "expected lifetime".

It is preferable that the above marker identification process be iterated one or more times by excluding one or more samples from the marker selection or ranking (*i.e.*, from the

calculation of correlation). Those samples being excluded are the ones that can not be predicted correctly from the previous iteration. Preferably, those samples excluded from marker selection in this iteration process are included in the classifier performance evaluation, to avoid overstating the performance.

5 5.1.3 CLASSIFIER GENESETS FOR PATIENT SUBSETS

This section provides exemplary sets of markers useful for the prognosis of breast cancer. The markers were identified according to the above methods in specific subsets of individuals with breast cancer. Generally, the marker sets can be used for prognosis of breast cancer patients that fall into five phenotypic categories based on criteria relevant to breast cancer prognosis, including estrogen receptor (ER) status, lymph node status, type of mutation(s) (*i.e.*, *BRCA1*-type or sporadic) and age at diagnosis. More specifically, patients, and tumors from which samples were taken, were classified as ER⁻, sporadic (*i.e.*, being both estrogen receptor negative and having a non-*BRCA1*-type tumor); ER⁻, *BRCA1* (*i.e.*, being both estrogen receptor negative and having a *BRCA1*-type tumor); ER⁺, ER/AGE high (*i.e.*, estrogen receptor positive with a high ratio of the log (ratio) of estrogen receptor gene expression to age); ER⁺, ER/AGE low, LN⁺ (*i.e.*, estrogen receptor positive with a low ratio of the log (ratio) of estrogen receptor gene expression to age, lymph node positive); and ER⁺, ER/AGE low, LN⁻ (*i.e.*, estrogen receptor positive with a low ratio of the log (ratio) of estrogen receptor gene expression to age, lymph node negative). The rationale for subdivision of the original patient set into these five subsets is detailed in the Examples (Section 6). The marker sets useful for each of the subsets above are provided in Tables 1-5, respectively.

Table 1: Geneset of 20 markers used to classify ER⁻, sporadic individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list
AF055033	IGFBP5	-2.12	0.88	0.54	insulin-like growth factor binding protein 5	Growth factor binding, Glycoprotein, Signal, 3D-structure
NM_000599	IGFBP5	-3.41	0.43	0.53	insulin-like growth factor binding protein 5	Growth factor binding, Glycoprotein, Signal, 3D-structure
L27560	IGFBP5	-4.55	0	0.52	EST	Hypothetical protein
AF052162	FLJ12443	-0.27	1.6	0.52	EST	Hypothetical protein
NM_001456	FLNA	-0.61	2.47	0.52	filamin A, alpha (actin binding protein 280)	Hypothetical protein, Actin-binding, Phosphorylation, Repeat,

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword_list
						Polymorphism, Disease mutation
NM_002205	ITGA5	-0.37	2.08	0.49	integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	Integrin, Cell adhesion, Receptor, Glycoprotein, Transmembrane, Signal, Calcium, Repeat
NM_013261	PPARGC1	0.09	1.54	0.47	peroxisome proliferative activated receptor, gamma, coactivator 1	
NM_001605	AARS	0.39	2.36	0.51	alanyl-tRNA synthetase	Aminoacyl-tRNA synthetase, Protein biosynthesis, Ligase, ATP-binding
X87949	HSPA5	-0.03	2.03	0.49	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	ATP-binding, Hypothetical protein, Endoplasmic reticulum, Signal
Contig50950_RC	NGEF	-1.17	3.2	0.52	neuronal guanine nucleotide exchange factor	
NM_005689	ABCB6	-0.51	2.26	0.48	ATP-binding cassette, sub-family B (MDR/TAP), member 6	ATP-binding, Transport, Transmembrane, Mitochondrion, Inner membrane, Transit peptide, Hypothetical protein
NM_004577	PSPH	-0.56	3.05	0.51	phosphoserine phosphatase	Hydrolase, Serine biosynthesis, Magnesium, Phosphorylation
NM_003832	PSPHL	-2.08	2.18	0.5	phosphoserine phosphatase-like	
NM_002422	MMP3	-0.96	2.54	0.5	matrix metalloproteinase 3 (stromelysin 1, progelatinase)	Hydrolase, Metalloprotease, Glycoprotein, Zinc, Zymogen, Calcium, Collagen degradation, Extracellular matrix, Signal, Polymorphism, 3D-structure
Contig37562_RC		-3.42	-6.02	-0.59	ESTs	
NM_018465	MDS030	-0.82	-3.28	-0.58	uncharacterized hematopoietic stem/progenitor cells protein MDS030	Hypothetical protein
Contig54661_RC		-0.79	-2.08	-0.54	ESTs	
AB032969	KIAA1143	-0.6	-2.85	-0.53	KIAA1143 protein	Hypothetical protein
Contig55353_RC	KIAA1915	-0.27	-1.82	-0.47	KIAA1915 protein	Hypothetical protein
NM_005213	CSTA	2.11	-3.4	-0.49	cystatin A (stefin A)	Thiol protease inhibitor, 3D-structure

Table 2. Geneset of 10 markers used to classify ER⁻, *BRCA1* individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_list
AF005487		6.08	0.5	-0.79	HLA-DRB6	Homo sapiens MHC class II antigen (DRB6)	MHC

						mRNA, HLA-DRB6*0201 allele, sequence.	
Contig50728_RC		4.02	0.25	-0.77		ESTs, Weakly similar to S26650 DNA-binding protein 5 - human [H.sapiens]	
Contig53598_RC		8.41	3.26	-0.77	FLJ11413	hypothetical protein FLJ11413	Hypothetical protein
NM_002888	RARR ES1	6.9	0.05	-0.87	RARR ES1	retinoic acid receptor responder (tazarotene induced) 1	Receptor, Transmembrane, Signal-anchor
NM_005218	DEFB1	5.14	-3.02	-0.81	DEFB1	defensin, beta 1	Antibiotic, Signal, 3D-structure
U17077	BENE	2.72	-1.72	-0.77	BENE	BENE protein	Transmembrane
Contig14683_RC		1.29	-2.31	-0.74		ESTs	
Contig53641_RC		-3.29	4.23	0.75	MAGE-E1	MAGE-E1 protein	Hypothetical protein
Contig56678_RC		-6.7	-9.73	-0.82		ESTs, Highly similar to THYA_HUMAN Prothymosin alpha [H.sapiens]	
NM_005461	KRML	0.88	-3.38	-0.75	MAFB	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	Transcription regulation, Repressor, DNA-binding, Nuclear protein, Hypothetical protein

Table 3. Geneset of 50 markers used to classify ER+, ER/AGE high individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list
NM_003600	STK15	-2.93	2.08	0.8	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase
NM_003158	STK6	-1.57	1.42	0.78	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase
NM_007019	UBCH10	-2.98	2.62	0.81	ubiquitin-conjugating enzyme E2C	Hypothetical protein, Ubl conjugation pathway, Ligase, Multigene family, Mitosis, Cell cycle, Cell division
NM_013277	ID-GAP	-2.43	2.43	0.77	Rac GTPase activating protein 1	Hypothetical protein
NM_004336	BUB1	-2.04	1.39	0.77	BUB1 budding uninhibited by benzimidazoles 1 homolog (yeast)	Transferase, Serine/threonine-protein kinase, ATP-binding, Cell cycle, Nuclear protein, Mitosis, Phosphorylation, Polymorphism
NM_006607	PTTG2	-1.71	1.49	0.72	pituitary tumor-transforming 2	
AK001166	FLJ11252	-1.33	0.99	0.71	hypothetical protein FLJ11252	Hypothetical protein
NM_004701	CCNB2	-4.62	2.01	0.81	cyclin B2	Cyclin, Cell cycle, Cell division, Mitosis
Contig57584_RC		-3.68	2.04	0.78	likely ortholog of mouse gene rich cluster, C8 gene	
NM_006845	KNSL6	-4.13	1.05	0.73	kinesin-like 6 (mitotic)	Hypothetical protein,

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list
					centromere-associated kinesin)	Motor protein, Microtubules, ATP-binding, Coiled coil, Nuclear protein
Contig38901_RC		-3.08	1.15	0.75	hypothetical protein MGC45866	Hypothetical protein
NM_018410	DKFZp762E1312	-4.38	1.49	0.75	hypothetical protein DKFZp762E1312	Hypothetical protein
NM_003981	PRC1	-3.52	2.17	0.78	protein regulator of cytokinesis 1	
NM_001809	CENPA	-5.04	0.98	0.75	centromere protein A, 17kDa	Hypothetical protein, Chromosomal protein, Nuclear protein, DNA-binding, Centromere, Antigen
NM_003504	CDC45L	-2.67	1.22	0.73	CDC45 cell division cycle 45-like (S. cerevisiae)	DNA replication, Cell cycle, Nuclear protein, Cell division
Contig41413_RC		-5.43	2.15	0.74	ribonucleotide reductase M2 polypeptide	Oxidoreductase, DNA replication, Iron
NM_004217	STK12	-2.17	0.73	0.72	serine/threonine kinase 12	Hypothetical protein, ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase
NM_002358	MAD2L1	-2.65	2.27	0.83	MAD2 mitotic arrest deficient-like 1 (yeast)	Cell cycle, Mitosis, Nuclear protein, 3D-structure
NM_014321	ORC6L	-2.73	1.8	0.75	origin recognition complex, subunit 6 homolog-like (yeast)	Hypothetical protein, DNA replication, Nuclear protein, DNA-binding
NM_012291	KIAA0165	-1.52	1.55	0.71	extra spindle poles like 1 (S. cerevisiae)	Hypothetical protein
NM_004203	PKMYT1	-3.64	2.2	0.7	retinoblastoma-like 2 (p130)	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase, Transcription regulation, DNA-binding, Nuclear protein, Cell cycle, Phosphorylation, Anti-oncogene
M96577	E2F1	-2.14	1.42	0.75	E2F transcription factor 1	Transcription regulation, Activator, DNA-binding, Nuclear protein, Phosphorylation, Cell cycle, Apoptosis, Polymorphism
NM_002266	KPNA2	-3.77	1.78	0.71	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Transport, Protein transport, Repeat, Nuclear protein, Polymorphism
Contig31288_RC		-2.63	0.7	0.68	ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens] [H.sapiens]	
NM_014501	E2-EPF	-1.55	1.93	0.7	ubiquitin carrier protein	Ubi conjugation pathway, Ligase, Multigene family

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list
NM_001168	BIRC5	-5.76	2.01	0.78	baculoviral IAP repeat-containing 5 (survivin)	Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D-structure, Hypothetical protein, Protease, Receptor
NM_003258	TK1	-4.57	1.38	0.71	thymidine kinase 1, soluble	Transferase, Kinase, DNA synthesis, ATP-binding
NM_001254	CDC6	-2.46	0.28	0.72	CDC6 cell division cycle 6 homolog (S. cerevisiae)	ATP-binding, Cell division
NM_004900	DJ742C19.2	-2.96	0.13	0.69	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B	Hydrolase
NM_004702	CCNE2	-3.12	2.13	0.81	cyclin E2	Cell cycle, Cell division, Cyclin, Hypothetical protein, Phosphorylation, Alternative splicing, Nuclear protein
AL160131		-3.07	2.42	0.7	hypothetical protein MGC861	Hypothetical protein
NM_016359	LOC51203	-3.22	2.61	0.76	nucleolar protein ANKT	Hypothetical protein, Nuclear protein
NM_004856	KNSL5	-1.52	1.1	0.71	kinesin-like 5 (mitotic kinesin-like protein 1)	Motor protein, Cell division, Microtubules, ATP-binding, Coiled coil, Mitosis, Cell cycle, Nuclear protein
NM_000057	BLM	-1.54	0.76	0.71	Bloom syndrome	Hydrolase, Helicase, ATP-binding, DNA-binding, Nuclear protein, DNA replication, Disease mutation
NM_018455	BM039	-2.44	1.18	0.7	uncharacterized bone marrow protein BM039	
NM_002106	H2AFZ	-2.49	1.53	0.72	H2A histone family, member Z	Chromosomal protein, Nucleosome core, Nuclear protein, DNA-binding, Multigene family
Contig64688		-2.68	3.1	0.73	hypothetical protein FLJ23468	Hypothetical protein
Contig44289_RC		-1.65	1.6	0.67	ESTs	
Contig28552_RC		-1.37	1.53	0.68	diaphanous homolog 3 (Drosophila)	Hypothetical protein, Coiled coil, Repeat, Alternative splicing
Contig46218_RC		-1.31	1.56	0.68	ESTs, Weakly similar to T19201 hypothetical protein C11G6.3 - Caenorhabditis elegans [C. elegans]	
Contig28947_RC		-1.3	0.98	0.67	cell division cycle 25A	Hypothetical protein, Cell division, Mitosis, Hydrolase, Alternative splicing, Multigene family, 3D-structure
NM_016095	LOC5165	-1.4	2.13	0.67	HSPC037 protein	Hypothetical protein

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Corre- lation	Description	Sp_xref_keyword_list
	9					
NM_003090	SNRPA1	-3.26	0.95	0.7	small nuclear ribonucleoprotein polypeptide A'	Hypothetical protein, Nuclear protein, RNA-binding, Ribonucleoprotein, Leucine-rich repeat, Repeat, 3D-structure
NM_002811	PSMD7	-2.48	1.89	0.7	proteasome (prosome, macropain) 26S subunit, non-ATPase, 7 (Mov34 homolog)	Proteasome
Contig38288_RC		-2.34	0.97	0.67	hypothetical protein DKFZp762A2013	Hypothetical protein
NM_003406	YWHAZ	-1.5	2.79	0.68	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide	Brain, Neurone, Phosphorylation, Acetylation, Multigene family, 3D-structure
AL137540	NTN4	2.13	-4.61	-0.69	netrin 4	Hypothetical protein, Laminin EGF-like domain, Signal
AL049367		1.9	-3.2	-0.68	EST	Transducer, Prenylation, Lipoprotein, Multigene family, Acetylation
NM_013409	FST	1.04	-5.78	-0.69	follistatin	Glycoprotein, Repeat, Signal, Alternative splicing
NM_000060	BTD	3.1	-1.45	-0.67	biotinidase	Hydrolase, Glycoprotein, Signal, Disease mutation

Table 4. Geneset of 50 markers used to classify ER+, ER/AGE low, LN+ individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword_list
NM_006417	MTAP44	-1.5	3	0.69	Fc fragment of IgG, low affinity IIb, receptor for (CD32)	Hydrolase, Hypothetical protein, Immunoglobulin domain, IgG-binding protein, Receptor, Transmembrane, Glycoprotein, Signal, Repeat, Multigene family, Polymorphism, NAD, One-carbon metabolism, Serine protease, Zymogen, Protease, Alternative splicing, Chromosomal translocation, Proto-oncogene, Galactin, Lectin, Antigen
NM_006820	GS3686	-4.3	4.06	0.69	chromosome 1 open reading frame 29	Hypothetical protein

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword_list
NM_001548	IFIT1	-3.4	4.27	0.71	Interferon-induced protein with tetratricopeptide repeats 1	Repeat, TPR repeat, Interferon induction
Contig41538_RC		-2.5	3.16	0.68	ESTs, Moderately similar to hypothetical protein FLJ20489 [<i>Homo sapiens</i>]	
NM_016816	OAS1	-1.7	3.29	0.75	2',5'-oligoadenylate synthetase 1, 40/46kDa	RNA-binding, Transferase, Nucleotidyltransferase, Interferon induction, Alternative splicing
Contig51660_RC		-2.1	2.65	0.66	28kD interferon responsive protein	Transmembrane
Contig43645_RC		-4.8	1.44	0.63	<i>Homo sapiens</i> , clone IMAGE:4428577, mRNA, partial cds	Hypothetical protein
AF026941		-4.6	2.71	0.63	EST, Weakly similar to 2004399A chromosomal protein [<i>Homo sapiens</i>]	Hypothetical protein
NM_007315	STAT1	-3.5	1.8	0.59	signal transducer and activator of transcription 1, 91kDa	Transcription regulation, DNA-binding, Nuclear protein, Phosphorylation, SH2 domain, Alternative splicing, 3D-structure
NM_002038	G1P3	-4.1	5.64	0.79	interferon, alpha-inducible protein (clone IFI-6-16)	Interferon induction, Transmembrane, Signal, Alternative splicing
NM_005101	ISG15	-5.6	5.34	0.77	interferon-stimulated protein, 15 kDa	Interferon induction, Repeat
NM_002462	MX1	-6.1	0.83	0.56	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)	Hypothetical protein, Interferon induction, GTP-binding, Multigene family, Antiviral
NM_005532	IFI27	-5.8	2.81	0.59	interferon, alpha-inducible protein 27	Interferon induction, Transmembrane
NM_002346	LY6E	-2.1	3.58	0.75	lymphocyte antigen 6 complex, locus E	Signal, Antigen, Multigene family, Membrane, GPI-anchor
NM_016817	OAS2	-3.6	1.89	0.59	2'-5'-oligoadenylate synthetase 2, 69/71kDa	RNA-binding, Transferase, Nucleotidyltransferase, Repeat, Interferon induction, Alternative splicing, Myristate
Contig44909_RC		-2.3	1.13	0.55	hypothetical protein BC012330	Hypothetical protein
NM_017414	USP18	-4.1	3.37	0.72	ubiquitin specific protease 18	Ubi conjugation pathway, Hydrolase, Thiol protease, Multigene family
NM_004029	IRF7	-2.4	3.67	0.66	interferon regulatory factor 7	Collagen, Transcription regulation, DNA-binding, Nuclear protein, Activator, Alternative splicing

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword_list
NM_004335	BST2	-3.2	3.22	0.57	bone marrow stromal cell antigen 2	Transmembrane, Glycoprotein, Signal- anchor, Polymorphism
NM_002759	PRKR	-2.4	1.8	0.58	protein kinase, interferon-inducible double stranded RNA dependent	Transferase, Serine/threonine-protein kinase, ATP-binding, Repeat, Phosphorylation, Interferon induction, RNA- binding, 3D-structure
NM_006332	IFI30	-3.8	2.65	0.64	interferon, gamma- inducible protein 30	Oxidoreductase, Interferon induction, Glycoprotein, Lysosome, Signal, Hypothetical protein
NM_009587	LGALS9	-3.2	2.08	0.6	lectin, galactoside- binding, soluble, 9 (galectin 9)	Galaptin, Lectin, Repeat, Alternative splicing
NM_003641	IFITM1	-2.4	5.54	0.63	interferon induced transmembrane protein 1 (9-27)	Interferon induction, Transmembrane
NM_017523	HSXIAPA F1	-1	2.84	0.7	XIAP associated factor-1	Hypothetical protein
NM_014314	RIG-I	-1.3	3.55	0.62	RNA helicase	ATP-binding, Helicase, Hydrolase, Hypothetical protein
Contig47563_RC		-2.2	3.11	0.56	ESTs	
AI497657_RC		-4.4	5.61	0.74	guanine nucleotide binding protein 4	Transducer, Prenylation, Lipoprotein, Multigene family
NM_000735	CGA	-4.3	2.5	0.58	glycoprotein hormones, alpha polypeptide	Hormone, Glycoprotein, Signal, 3D-structure
NM_004988	MAGEA1	-1.4	6.31	0.64	melanoma antigen, family A, 1 (directs expression of antigen MZ2-E)	Antigen, Multigene family, Polymorphism, Tumor antigen
Contig54242_RC		-1.2	4.1	0.65	chromosome 17 open reading frame 26	Hypothetical protein
NM_004710	SYNGR2	-1.4	3.01	0.54	synaptogyrin 2	Transmembrane
NM_001168	BIRC5	-3.7	3.39	0.64	baculoviral IAP repeat-containing 5 (survivin)	Hypothetical protein, Protease, Receptor, Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D-structure
Contig41413_RC		-4.4	2.61	0.57	ribonucleotide reductase M2 polypeptide	Oxidoreductase, DNA replication, Iron
NM_004203	PKMYT1	-3.4	3.79	0.6	retinoblastoma-like 2 (p130)	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase, Transcription regulation, DNA-binding, Nuclear protein, Cell cycle, Phosphorylation, Anti- oncogene

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Description	Sp_xref_keyword_list
Contig48913_RC		-3.1	1.72	0.55	<i>Homo sapiens</i> , Similar to hypothetical protein PRO1722, clone MGC:15692 IMAGE:3351479, mRNA, complete cds	
NM_005804	DDXL	-2.5	1.42	0.58	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 39	ATP-binding, Helicase, Hydrolase, Hypothetical protein
NM_016359	LOC51203	-1.7	3.6	0.57	nucleolar protein ANKT	Hypothetical protein, Nuclear protein
NM_001645	APOC1	-2.9	3.43	0.58	apolipoprotein C-I	Plasma, Lipid transport, VLDL, Signal, 3D-structure, Polymorphism
Contig37895_RC		-2	2.05	0.55	ESTs	
NM_005749	TOB1	-1.3	4.96	0.59	transducer of ERBB2, 1	Phosphorylation
NM_000269	NME1	-1.3	2.98	0.55	non-metastatic cells 1, protein (NM23A) expressed in	Transferase, Kinase, ATP-binding, Nuclear protein, Anti-oncogene, Disease mutation
NM_014462	LSM1	-1	4.5	0.57	Lsm1 protein	Nuclear protein, Ribonucleoprotein, mRNA splicing, mRNA processing, RNA-binding
Contig31221_RC		-1.4	3.83	0.56	HTPAP protein	
NM_005326	HAGH	-1.9	4.29	0.57	hydroxyacyl glutathione hydrolase	Hydrolase, Zinc, 3D-structure
Contig42342_RC		0.78	-3.2	-0.6	<i>Homo sapiens</i> cDNA FLJ39417 fis, clone PLACE6016942	Hypothetical protein
AL137540	NTN4	2.24	-3.9	-0.6	netrin 4	Laminin EGF-like domain, Signal, Hypothetical protein
Contig40434_RC		1.64	-5.6	-0.6	wingless-type MMTV integration site family, member 5A	Developmental protein, Glycoprotein, Signal
Contig1632_RC		1.03	-3.9	-0.6	hypothetical protein MGC17921	Hypothetical protein
NM_014246	CELSR1	0.95	-4.6	-0.6	cadherin, EGF LAG seven-pass G-type receptor 1 (flamingo homolog, <i>Drosophila</i>)	G-protein coupled receptor, Transmembrane, Glycoprotein, EGF-like domain, Calcium-binding, Laminin EGF-like domain, Repeat, Developmental protein, Hydroxylation, Signal, Alternative splicing, Hypothetical protein
NM_005139	ANXA3	1.26	-6.2	-0.6	annexin A3	Annexin, Calcium/phospholipid-binding, Repeat, Phospholipase A2 inhibitor, 3D-structure, Polymorphism

Table 5. Geneset of 65 markers used to classify ER+, ER/AGE low, LN⁻ individuals.

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_li st
M55914	MPB1	-2.82	1.25	0.5	ENO1	enolase 1, (alpha)	DNA-binding, Transcription regulation, Repressor, Nuclear protein, Lyase, Glycolysis, Magnesium, Multigene family, Hypothetical protein
NM_005945	MPB1	-3.06	1.19	0.49	ENO1	Homo sapiens enolase 1, (alpha) (ENO1), mRNA.	Glycolysis, Hypothetical protein, Lyase, Magnesium, DNA-binding, Transcription regulation, Repressor, Nuclear protein, Multigene family
NM_001428	ENO1	-2.53	1.18	0.46	ENO1	enolase 1, (alpha)	DNA-binding, Transcription regulation, Repressor, Nuclear protein, Lyase, Glycolysis, Magnesium, Multigene family, Hypothetical protein
NM_001216	CA9	-4.72	1.49	0.6	CA9	carbonic anhydrase IX	Lyase, Zinc, Transmembrane, Glycoprotein, Antigen, Signal, Nuclear protein, Polymorphism
NM_001124	ADM	-5.68	2.99	0.56	ADM	adrenomedullin	Hormone, Amidation, Cleavage on pair of basic residues, Signal
NM_000584	IL8	-2.45	2.04	0.54	IL8	interleukin 8	Cytokine, Chemotaxis, Inflammatory response, Signal, Alternative splicing, 3D-structure
D25328	PFKP	-4.19	3.29	0.56	PFKP	Phosphofructo- kinase, platelet	Kinase, Transferase, Glycolysis, Repeat, Allosteric enzyme, Phosphorylation, Magnesium, Multigene family
NM_006096	NDRG1	-5.45	5.97	0.77	NDRG1	N-myc downstream regulated gene 1	Hypothetical protein, Nuclear protein, Repeat
NM_004994	MMP9	-5.53	1.07	0.49	MMP9	matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	Hydrolase, Metalloprotease, Glycoprotein, Zinc, Zymogen, Calcium, Collagen degradation, Extracellular matrix, Repeat, Signal,

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_li st
							Polymorphism, 3D- structure
NM_003311	TSSC3	-4.57	5.58	0.68	TSSC3	tumor suppressing subtransferable candidate 3	
NM_006086	TUBB4	-5.19	2.85	0.59	TUBB4	tubulin, beta, 4	G-protein coupled receptor, Transmembrane, Glycoprotein, Phosphorylation, Lipoprotein, Palmitate, Polymorphism, Hypothetical protein, GTP-binding, Receptor, Microtubules, Multigene family
NM_006115	PRAME	-4.48	2.77	0.61	PRAME	preferentially expressed antigen in melanoma	Antigen
NM_004345	CAMP	-2.02	1.37	0.49	CAMP	cathelicidin antimicrobial peptide	Antibiotic, Signal
NM_018455	BM039	-2.34	0.76	0.47	BM039	uncharacterized bone marrow protein BM039	
Contig49169_RC		-1.17	1.5	0.46	SUV39H2	suppressor of variegation 3-9 (Drosophila) homolog 2; hypothetical protein FLJ23414	Hypothetical protein, Nuclear protein
Contig45032_RC		-1.37	0.77	0.45	FLJ14813	hypothetical protein FLJ14813	Hypothetical protein, ATP-binding, Kinase, Serine/threonine- protein kinase, Transferase
NM_000917	P4HA1	-1.54	4.31	0.62	P4HA1	procollagen- proline, 2- oxoglutarate 4- dioxygenase (proline 4- hydroxylase), alpha polypeptide I	Dioxygenase, Collagen, Oxidoreductase, Iron, Vitamin C, Alternative splicing, Glycoprotein, Endoplasmic reticulum, Signal
NM_002046	GAPD	-2.51	3.42	0.6	GAPD	glyceraldehyde-3- phosphate dehydrogenase	Glycolysis, NAD, Oxidoreductase, Hypothetical protein, Multigene family
NM_000365	TPI1	-1.81	2.94	0.56	TPI1	triosephosphate isomerase 1	Fatty acid biosynthesis, Gluconeogenesis, Glycolysis, Isomerase, Pentose shunt, Disease mutation, Polymorphism, 3D-

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_li st
							structure
NM_014364	GAPDS	-1.08	2.88	0.58	GAPDS	glyceraldehyde-3-phosphate dehydrogenase, testis-specific	Glycolysis, Oxidoreductase, NAD
NM_005566	LDHA	-2.01	4.01	0.59	LDHA	lactate dehydrogenase A	Oxidoreductase, NAD, Glycolysis, Multigene family, Disease mutation, Polymorphism
NM_000291	PGK1	-2.28	1.68	0.51	PGK1	phosphoglycerate kinase 1	Kinase, Transferase, Multigene family, Glycolysis, Acetylation, Disease mutation, Polymorphism, Hereditary hemolytic anemia
NM_016185	LOC51155	-2.33	2.82	0.59	HN1	hematological and neurological expressed 1	
NM_001168	BIRC5	-4.33	2.78	0.55	BIRC5	baculoviral IAP repeat-containing 5 (survivin)	Apoptosis, Thiol protease inhibitor, Alternative splicing, 3D-structure, Hypothetical protein, Protease, Receptor
NM_002266	KPNA2	-3.75	1.34	0.47	KPNA2	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Transport, Protein transport, Repeat, Nuclear protein, Polymorphism
Contig31288_RC		-2.1	1.27	0.5		ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens] [H.sapiens]	
NM_000269	NME1	-2.15	3.43	0.55	NME1	non-metastatic cells 1, protein (NM23A) expressed in	Transferase, Kinase, ATP-binding, Nuclear protein, Anti-oncogene, Disease mutation
NM_003158	STK6	-1.23	1.73	0.45	STK6	serine/threonine kinase 6	ATP-binding, Kinase, Serine/threonine-protein kinase, Transferase
NM_007274	HBACH	-1.83	2.73	0.51	BACH	brain acyl-CoA hydrolase	Hydrolase, Serine esterase, Repeat
Contig55188_RC		-2.36	3.28	0.47	FLJ22341	hypothetical protein FLJ22341	Hypothetical protein
NM_002061	GCLM	-1.06	1.76	0.48	GCLM	glutamate-cysteine ligase, modifier subunit	Ligase, Glutathione biosynthesis
NM_004207	SLC16A3	-3.11	5.07	0.67	SLC16A3	solute carrier family 16 (monocarboxylic acid transporters), member 3	Transport, Symport, Transmembrane, Multigene family
NM_000582	SPP1	-5.09	5.47	0.53	SPP1	secreted phosphoprotein 1	Hypothetical protein, Glycoprotein, Sialic

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_l ist
						(osteopontin, bone sialoprotein I, early T-lymphocyte activation 1)	acid, Biomineralization, Cell adhesion, Phosphorylation, Signal, Alternative splicing
NM_001109	ADAM8	-2.5	3.74	0.45	ADAM8	a disintegrin and metalloproteinase domain 8	Hydrolase, Metalloprotease, Zinc, Signal, Glycoprotein, Transmembrane, Antigen
D50402	SLC11A1	-1.05	3.46	0.53	SLC11A1	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1	Transport, Iron transport, Transmembrane, Glycoprotein, Macrophage, Polymorphism
AL080235	DKFZP586E1621	-1.23	1.96	0.51	RIS1	Ras-induced senescence 1	Hypothetical protein
Contig40552_RC		-1.26	3.96	0.54	FLJ25348	hypothetical protein FLJ25348	Hypothetical protein
Contig52490_RC		-0.64	3.33	0.61	LOC116238	hypothetical protein BC014072	
NM_006461	DEEPEST	-2.1	1.85	0.46	SPAG5	sperm associated antigen 5	Hypothetical protein
Contig56503_RC		-4.3	3.39	0.55	MGC9753	hypothetical gene MGC9753	Hypothetical protein
Contig63525		-1.91	3.34	0.5	FLJ13352	hypothetical protein FLJ13352	Hypothetical protein
NM_001909	CTSD	-0.83	4.6	0.51	CTSD	cathepsin D (lysosomal aspartyl protease)	Hydrolase, Aspartyl protease, Glycoprotein, Lysosome, Signal, Zymogen, Polymorphism, Alzheimer's disease, 3D-structure
NM_005063	SCD	-2.57	5.15	0.48	SCD	stearoyl-CoA desaturase (delta-9-desaturase)	Hypothetical protein, Endoplasmic reticulum, Fatty acid biosynthesis, Iron, Oxidoreductase, Transmembrane
NM_005165	ALDOC	-2.43	5.02	0.48	ALDOC	aldolase C, fructose-bisphosphate	Lyase, Schiff base, Glycolysis, Multigene family
NM_000363	TNNI3	-0.54	3.58	0.48	TNNI3	troponin I, cardiac	Hypothetical protein, Muscle protein, Actin-binding, Acetylation, Disease mutation, Cardiomyopathy, Receptor, Signal
AF035284		-1.63	3.28	0.47	FADS1	EST	Heme, Hypothetical protein
Contig30875_RC		-0.88	3	0.6		ESTs	

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_li st
NM_018487	HCA112	-0.7	3.54	0.58	HCA112	hepatocellular carcinoma-associated antigen 112	Hypothetical protein
NM_001323	CST6	-1.63	3.84	0.57	CST6	cystatin E/M	Thiol protease inhibitor, Signal, Glycoprotein
NM_006516	SLC2A1	-1.66	2.22	0.46	SLC2A1	solute carrier family 2 (facilitated glucose transporter), member 1	Transmembrane, Sugar transport, Transport, Glycoprotein, Multigene family, Disease mutation
NM_007267	LAK-4P	-1.04	3.28	0.61	EVIN1	expressed in activated T/LAK lymphocytes	Hypothetical protein
NM_004710	SYNGR2	-0.84	4.81	0.56	SYNGR2	synaptogyrin 2	Transmembrane
Contig63649_RC		-1.34	6.3	0.75		ESTs, Weakly similar to 2004399A chromosomal protein [Homo sapiens] [H.sapiens]	
NM_003376	VEGF	-2.12	2.42	0.46	VEGF	vascular endothelial growth factor	Hypothetical protein, Mitogen, Angiogenesis, Growth factor, Glycoprotein, Signal, Heparin-binding, Alternative splicing, Multigene family, 3D-structure
NM_000799	EPO	-0.75	4.01	0.69	EPO	erythropoietin	Erythrocyte maturation, Glycoprotein, Hormone, Signal, Pharmaceutical, 3D-structure
NM_006014	DXS9879E	-1.85	3.44	0.54	DXS9879E	DNA segment on chromosome X (unique) 9879 expressed sequence	
NM_007183	PKP3	-0.91	4.14	0.48	PKP3	plakophilin 3	Cell adhesion, Cytoskeleton, Structural protein, Nuclear protein, Repeat
D13642	SF3B3	-0.65	2.28	0.48	SF3B3	splicing factor 3b, subunit 3, 130kDa	Hypothetical protein, Spliceosome, mRNA processing, mRNA splicing, Nuclear protein
NM_003756	EIF3S3	-1.85	2.19	0.46	EIF3S3	eukaryotic translation initiation factor 3, subunit 3 gamma, 40kDa	Initiation factor, Protein biosynthesis
Contig47096_RC		-0.41	4.52	0.54	PFKFB4	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4	Kinase, Multifunctional enzyme, Transferase, Hydrolase, ATP-binding,

Accession/ Contig No.	Gene	Avg good xdev	Avg poor xdev	Correl- ation	Sequence name	Description	Sp_xref_keyword_li st
							Phosphorylation, Multigene family
NM_004209	SYNGR 3	-0.31	3.67	0.53	SYNGR3	synaptogyrin 3	Transmembrane
Contig3464_RC		0.99	-5.81	-0.52		ESTs	
Contig31646_RC		1.1	-7.76	-0.5	COL14A1	collagen, type XIV, alpha 1 (undulin)	Extracellular matrix, Glycoprotein, Hypothetical protein, Collagen, Signal
Contig49388_RC		1.73	-1.75	-0.51	FLJ13322	hypothetical protein FLJ13322	Hypothetical protein
Contig41887_RC		0.37	-5.74	-0.47	LOC12422 0	similar to common salivary protein 1	Hypothetical protein

Table 8 Good and poor prognosis templates: mean subtracted log(intensity) values for each of the seventy markers for 44 breast cancer patients having a good prognosis (C1) or 34 breast cancer patients having a poor prognosis (C2) (see, PCT publication WO 2004/065545, published August 5, 2004, which is incorporated herein by reference in its entirety)

Marker Accession #	C1 (good prognosis template)	C2 (poor prognosis template)	SEQ ID NO
AL080059	-5.161569	0.043019	367
Contig63649_RC	-1.440895	0.966702	361
Contig46218_RC	-0.937662	0.815081	321
NM_016359	-1.49878	0.872829	253
AA555029_RC	-1.283504	0.543442	368
NM_003748	1.355486	-0.254201	369
Contig38288_RC	-1.237495	1.085461	297
NM_003862	0.981236	-1.619658	370
Contig28552_RC	-1.296043	1.067545	281
Contig32125_RC	0.855155	-0.7338	371
U82987	1.256206	-1.362807	372
AL137718	-0.55046	0.68754	373
AB037863	0.819061	-1.621057	374
NM_020188	-1.137582	0.673123	375
NM_020974	-0.463953	-5.623268	376
NM_000127	-0.618568	0.552726	377
NM_002019	-1.409168	0.547285	378
NM_002073	-1.577177	0.417352	379
NM_000436	-0.722574	0.599239	380
NM_004994	-5.561089	-2.180659	165
Contig55377_RC	0.805683	-1.16728	381
Contig35251_RC	-0.931146	0.607562	382
Contig25991	-0.720727	1.045949	383
NM_003875	-1.206839	1.163244	384
NM_006101	-0.879965	0.628296	385
NM_003882	0.529121	-0.467098	386
NM_003607	-0.959094	0.709653	387

Marker Accession #	C1 (good prognosis template)	C2 (poor prognosis template)	SEQ ID NO
AF073519	-1.451486	0.163988	388
AF052162	-1.145575	0.192391	9
NM_000849	0.944742	-1.499473	389
Contig32185_RC	-0.887643	0.688257	390
NM_016577	-2.762008	0.081637	391
Contig48328_RC	0.405401	-2.946904	392
Contig46223_RC	0.805424	-0.581849	393
NM_015984	-1.056531	0.522176	394
NM_006117	1.129928	-1.262974	395
AK000745	-2.475715	-0.013002	396
Contig40831_RC	-1.17091	0.435754	397
NM_003239	0.457773	-2.150499	398
NM_014791	-1.14862	0.383018	399
X05610	-0.768514	0.637938	400
NM_016448	-0.713264	0.632638	401
NM_018401	0.618921	-0.286778	402
NM_000788	-0.995116	0.50246	403
Contig51464_RC	-0.663538	0.765975	404
AL080079	-1.794821	0.43708	405
NM_006931	-0.846271	0.915602	406
AF257175	1.122354	-0.721924	407
NM_014321	-1.820261	0.482287	241
NM_002916	-0.966852	0.599925	408
Contig55725_RC	-2.935162	0.623397	409
Contig24252_RC	-2.004671	0.263597	410
AF201951	0.355839	-2.296556	411
NM_005915	-0.586121	0.827714	412
NM_001282	0.762645	-0.970418	413
Contig56457_RC	-0.920808	0.588269	414
NM_000599	-3.612469	-0.714313	51
NM_020386	-0.46073	0.699313	415
NM_014889	-1.678462	0.2362	416
AF055033	-2.505271	-0.07576	11
NM_006681	-0.631302	0.584119	417
NM_007203	-1.426446	0.504624	418
Contig63102_RC	0.521511	-1.266163	419
NM_003981	-2.521877	0.552669	420
Contig20217_RC	-0.363574	0.449022	421
NM_001809	-2.171301	0.328419	81
Contig2399_RC	-1.174844	0.602523	422
NM_004702	-1.560133	0.619078	155
NM_007036	-0.950633	0.34945	423
NM_018354	-1.392354	0.347831	424

5.2. DIAGNOSTIC AND PROGNOSTIC METHODS

5.2.1 SAMPLE COLLECTION

In the present invention, markers, such as target polynucleotide molecules or proteins, are extracted from a sample taken from an individual afflicted with a condition such as breast cancer. The sample may be collected in any clinically acceptable manner, but must be collected such that marker-derived polynucleotides (*i.e.*, RNA) are preserved (if gene
5 expression is to be measured) or proteins are preserved (if encoded proteins are to be measured). For example, mRNA or nucleic acids derived therefrom (*i.e.*, cDNA or amplified DNA) are preferably labeled distinguishably from standard or control polynucleotide molecules, and both are simultaneously or independently hybridized to a microarray comprising some or all of the markers or marker sets or subsets described above.

10 Alternatively, mRNA or nucleic acids derived therefrom may be labeled with the same label as the standard or control polynucleotide molecules, wherein the intensity of hybridization of each at a particular probe is compared. A sample may comprise any clinically relevant tissue sample, such as a tumor biopsy or fine needle aspirate, or a sample of bodily fluid, such as blood, plasma, serum, lymph, ascitic fluid, cystic fluid, urine or nipple exudate. The sample
15 may be taken from a human, or, in a veterinary context, from non-human animals such as ruminants, horses, swine or sheep, or from domestic companion animals such as felines and canines.

Methods for preparing total and poly(A)+ RNA are well known and are described generally in Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.),
20 Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)) and Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994)).

RNA may be isolated from eukaryotic cells by procedures that involve lysis of the cells and denaturation of the proteins contained therein. Cells of interest include wild-type
25 cells (*i.e.*, non-cancerous), drug-exposed wild-type cells, tumor- or tumor-derived cells, modified cells, normal or tumor cell line cells, and drug-exposed modified cells. Preferably, the cells are breast cancer tumor cells.

Additional steps may be employed to remove DNA. Cell lysis may be accomplished with a nonionic detergent, followed by microcentrifugation to remove the nuclei and hence
30 the bulk of the cellular DNA. In one embodiment, RNA is extracted from cells of the various types of interest using guanidinium thiocyanate lysis followed by CsCl centrifugation to separate the RNA from DNA (Chirgwin *et al.*, *Biochemistry* 18:5294-5299 (1979)).

Poly(A)+ RNA is selected by selection with oligo-dT cellulose (*see* Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989). Alternatively, separation of RNA from DNA can be accomplished by organic extraction, for example, with hot phenol or
5 phenol/chloroform/isoamyl alcohol.

If desired, RNase inhibitors may be added to the lysis buffer. Likewise, for certain cell types, it may be desirable to add a protein denaturation/digestion step to the protocol.

For many applications, it is desirable to preferentially enrich mRNA with respect to other cellular RNAs, such as transfer RNA (tRNA) and ribosomal RNA (rRNA). Most
10 mRNAs contain a poly(A) tail at their 3' end. This allows them to be enriched by affinity chromatography, for example, using oligo(dT) or poly(U) coupled to a solid support, such as cellulose or Sephadex™ (*see* Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994). Once bound, poly(A)+ mRNA is eluted from the affinity column using 2 mM EDTA/0.1% SDS.

15 The sample of RNA can comprise a plurality of different mRNA molecules, each different mRNA molecule having a different nucleotide sequence. In a specific embodiment, the mRNA molecules in the RNA sample comprise at least 5, 10, 15, 20, 25, 30, 40 or 50 different nucleotide sequences. More preferably, the mRNA molecules of the RNA sample comprise mRNA molecules corresponding to each of the marker genes. In another specific
20 embodiment, the RNA sample is a mammalian RNA sample.

In a specific embodiment, total RNA or mRNA from cells is used in the methods of the invention. The source of the RNA can be cells of a plant or animal, human, mammal, primate, non-human animal, dog, cat, mouse, rat, bird, yeast, eukaryote, prokaryote, etc. In specific embodiments, the method of the invention is used with a sample containing total
25 mRNA or total RNA from 1×10^6 cells or less. In another embodiment, proteins can be isolated from the foregoing sources, by methods known in the art, for use in expression analysis at the protein level.

Probes to the homologs of the marker sequences disclosed herein can be employed preferably when non-human nucleic acid is being assayed.

The prognosis may be carried out using a cellular constituent profile of any type of molecular markers. For example, sets of proteins informative for breast cancer prognosis may be used. Such protein markers may be measured, for example, by use of gel electrophoresis, such as one-dimensional polyacrylamide gel electrophoresis, two-dimensional polyacrylamide gel electrophoresis, nondenaturing polyacrylamide gel electrophoresis; isoelectric focusing gels, etc., by use of antibody arrays, etc. Of course, the particular template(s) used to classify the individual depends upon the type(s) of cellular constituents used as markers. For example, where nucleic acids (e.g., genes or nucleic acids derived from expressed genes) are used as markers, the template comprises nucleic acids (or the level of expression or abundance thereof); where proteins are used as markers, the template comprises proteins, for example, the level or abundance of those proteins; etc.

5.2.2. USE OF PROGNOSTIC GENESETS FOR BREAST CANCER

According to the present invention, once genesets informative for a plurality of subsets of a condition are identified, an individual is classified into one of these subsets and a prognosis is made based on the expression of the genes, or their encoded proteins, in the geneset for that subset in a breast cancer tumor sample taken from the individual. The patient's responsiveness to chemotherapy is then determined based on the patient's subset classification and the prognosis. Various methods are known in the art that can be used for classifying an expression profile. Exemplary methods are described in Section 5.2.4, *infra*.

For example, a particular hypothetical condition has three relevant phenotypic characteristics, A, B and C. In this example, based on these characteristics, genesets informative for prognosis of four patient subsets A^+B^+ ; $A^+B^-C^+$; $A^+B^-C^-$; and A^- are identified by the method described above. Thus, an individual having the condition would first be classified according to phenotypes A-C into one of the four patient subsets. In one embodiment, therefore, the invention provides for the classification of an individual having a condition into one of a plurality of patient subsets, wherein a set of genes informative for prognosis for the subset has been identified. A sample is then taken from the individual, and the levels of expression of the prognostically-informative genes in the sample is analyzed. In one embodiment, the expression level of each gene can be compared to the expression level of the corresponding gene in a control or reference sample to determine a differential expression level. The expression profile comprising expression levels, e.g., differential expression levels, of the plurality of genes is then compared to a template profile. In various

embodiments, the template profile is a good prognosis template comprising the average expression of informative genes in samples taken from good prognosis individuals classifiable into that patient subset; or a poor prognosis template comprising the average expression of informative genes in samples taken from poor prognosis individuals classifiable
5 into that patient subset; or a good prognosis profile comprising a set of mathematical values that represent gene expression levels of good prognosis individuals classifiable into that patient subset; etc.

In a specific embodiment, the phenotypic, genotypic and/or clinical classes for breast cancer are: ER⁻, *BRCA1* individuals; ER⁻, sporadic individuals; ER⁺, ER/AGE high
10 individuals; ER⁺, ER/AGE low, LN⁺ individuals; and ER⁺, ER/AGE low, LN⁻ individuals. In this embodiment, an individual may be classified as ER⁺ or ER⁻. If the individual is ER⁻, the individual is additionally classified as having a *BRCA1*-type or sporadic tumor. ER⁻ individuals are thus classified as ER⁻, *BRCA1* or ER⁻, sporadic. Alternatively, if the individual is classified as ER⁺, the individual is additionally classified as having a high or
15 low ratio of the log (ratio) of the level of expression of the gene encoding the estrogen receptor to the individual's age. Individuals having a low ratio are additionally classified as LN⁺ or LN⁻. ER⁺ individuals are thus classified as ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺, or ER⁺, ER/AGE low, LN⁻. Of course, the individual's ER status, tumor type, age and
20 LN status may be identified in any order, as long as the individual is classified into one of these five subsets.

In one embodiment, a breast cancer patient is first classified into one of the following patient subsets: ER⁻, *BRCA1* individuals; ER⁻, sporadic individuals; ER⁺, ER/AGE high individuals; ER⁺, ER/AGE low, LN⁺ individuals; and ER⁺, ER/AGE low, LN⁻ individuals. The patient is then classified as having a good prognosis or a poor prognosis by comparing
25 the patient's expression profile of a plurality of genes or their encoded proteins to a good and/or poor prognosis template profile of expression levels of said plurality of genes or their encoded proteins for the appropriate subset, and classifying said patient as having a good prognosis or poor prognosis if said patient expression profile has a high similarity to a good prognosis template or a poor prognosis template, respectively. In a preferred embodiment,
30 the good prognosis template for a patient subset comprises the average level of expression of each of said plurality of genes informative for prognosis in tumor samples from individuals classified into said subset who have a good prognosis or good outcome, while the poor

prognosis template comprises the average level of expression of each of said plurality of genes informative for prognosis in tumor samples from individuals classified into said subset who have a poor prognosis or poor outcome. In another specific embodiment, said good or poor prognosis template is a set of mathematical values representing the average level of expression of genes informative for prognosis in tumor samples of individuals classifiable into said subset who have a good prognosis, or who have a poor prognosis, respectively.

It is evident that the different patient subsets described herein reflect different molecular mechanisms of the initiation of tumor formation and metastasis. Thus, the genesets listed in tables 1-5 are also useful for diagnosing a person as having a particular type of breast cancer in the first instance. Thus, the invention also provides a method of diagnosing an individual as having a particular subtype of breast cancer, comprising determining an expression profile of a plurality of the genes for which markers are listed in Tables 1-5 in a sample from said individual; and comparing said expression profile to a template profile, where said template is representative of the expression of said plurality of genes in a breast cancer sample of said subtype of cancer, and on the basis of said comparison, diagnosing the individual as having said subtype of breast cancer. In a specific embodiment, said subtype of cancer is selected from the group consisting of ER⁻, *BRCA1* type; ER⁻, sporadic type; ER⁺, ER/AGE high type; ER⁺, ER/AGE low, LN⁺ type; and ER/AGE low, LN⁻ type. In another specific embodiment, said template comprises the average levels of expression of a plurality of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5. In another specific example, said comparing comprises determining the similarity of the expression profile of the genes for which markers are listed in each of Tables 1-5 in said sample taken from said individual to a template profile comprising levels of expression of the same genes for each of Tables 1-5, and determining whether the pattern of expression of said genes in said sample is most similar to the pattern of expression of the genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5.

In another embodiment, a breast cancer patient is classified as having a good prognosis or a poor prognosis by a method comprising: (a) classifying said patient as ER⁻, *BRCA1*; ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (b) determining an expression profile of a first plurality of genes in a cell sample taken from the patient, said first plurality of genes comprising at least two of the genes

corresponding to the markers Table 1 if said patient is classified as ER⁻, *BRCAl*; Table 2 if said patient is classified as ER⁻, sporadic; Table 3 if said patient is classified as ER⁺, ER/AGE high; Table 4 if said patient is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said patient is classified as ER⁺, ER/AGE low, LN⁻, wherein said patient is “ER/AGE high” if the ER level of the patient exceeds a predetermined value, and “ER/AGE low” if the ER level of the patient does not exceed said predetermined value. In a specific embodiment of this method, said predetermined value of ER is calculated as $ER = 0.1(AGE - 42.5)$, wherein AGE is the age of said patient. In another specific embodiment, said patient is ER⁻, *BRCAl*, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 1. In another specific embodiment, said patient is ER⁻, sporadic, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 2. In another specific embodiment, said patient is ER⁺, ER/AGE high, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 3. In another specific embodiment, said patient is ER⁺, ER/AGE low, LN⁺, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 4. In another specific embodiment, said patient is ER⁺, ER/AGE low, LN⁻, and said plurality of genes comprises at least 1, 2, 3, 4, 5, 10 or all of the genes for which markers are listed in Table 5. In another specific embodiment, the method additionally comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis.

Where information is available regarding the LN status of a breast cancer patient, the patient may be identified as having a “very good prognosis,” an “intermediate prognosis,” or a poor prognosis, which enables the refinement of treatment. In one embodiment, the invention provides a method of assigning a therapeutic regimen to a breast cancer patient, comprising: (a) classifying said patient as having a “poor prognosis,” “intermediate prognosis,” or “very good prognosis” on the basis of the levels of expression of at least five genes for which markers are listed in Table 1, Table 2, Table 3, Table 4 or Table 5; and (b) assigning said patient a therapeutic regimen, said therapeutic regimen (i) comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or (ii) comprising chemotherapy if said patient has any other combination of lymph node status and expression profile.

In another embodiment, a breast cancer patient is assigned a prognosis by a method comprising (a) determining the breast cancer patient's age, ER status, LN status and tumor type; (b) classifying said patient as ER⁻, sporadic; ER⁻, *BRCA1*; ER+, ER/AGE high; ER+, ER/AGE low, LN+; or ER+, ER/AGE low, LN⁻; (c) determining an expression profile comprising at least five genes in a cell sample taken from said breast cancer patient wherein markers for said at least five genes are listed in Table 1 if said patient is classified as ER⁻, sporadic; Table 2 if said patient is classified as ER⁻, *BRCA1*; Table 3 if said patient is classified as ER+, ER/AGE high; Table 4 if said patient is classified as ER+, ER/AGE low, LN+; or Table 5 if said patient is classified as ER+, ER/AGE high, LN⁻; (d) determining the similarity of the expression profile of said at least five genes to a template profile comprising levels of expression of said at least five genes to obtain a patient similarity value; (e) comparing said patient similarity value to selected first and second threshold values of similarity, respectively, wherein said second similarity threshold indicates greater similarity to said template expression profile than does said first similarity threshold; and (f) classifying said breast cancer patient as having a first prognosis if said patient similarity value exceeds said second threshold similarity values, a second prognosis if said patient similarity value exceeds said first threshold similarity value but does not exceed said second threshold similarity value, and a third prognosis if said patient similarity value does not exceed said first threshold similarity value. In a specific embodiment of the method, said first prognosis is a "very good prognosis," said second prognosis is an "intermediate prognosis," and said third prognosis is a "poor prognosis," wherein said breast cancer patient is assigned a therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or comprising chemotherapy if said patient has any other combination of lymph node status and expression profile.

The invention also provides a method of assigning a therapeutic regimen to a breast cancer patient, comprising: (a) determining the lymph node status for said patient; (b) determining the expression of at least five genes for which markers are listed in Table 5 in a cell sample from said patient, thereby generating an expression profile; (c) classifying said patient as having a "poor prognosis," "intermediate prognosis," or "very good prognosis" on the basis of said expression profile; and (d) assigning said patient a therapeutic regimen, said therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or

comprising chemotherapy if said patient has any other combination of lymph node status and classification. In a specific embodiment of this method, said therapeutic regimen assigned to lymph node negative patients classified as having an "intermediate prognosis" additionally comprises adjuvant hormonal therapy. In another specific embodiment of this method, said

5 classifying step (c) is carried out by a method comprising: (a) rank ordering in descending order a plurality of breast cancer tumor samples that compose a pool of breast cancer tumor samples by the degree of similarity between the expression profile of said at least five genes in each of said tumor samples and the expression profile of said at least five genes across all remaining tumor samples that compose said pool, said degree of similarity being expressed as

10 a similarity value; (b) determining an acceptable number of false negatives in said classifying step, wherein a false negative is a breast cancer patient for whom the expression levels of said at least five genes in said cell sample predicts that said breast cancer patient will have no distant metastases within the first five years after initial diagnosis, but who has had a distant metastasis within the first five years after initial diagnosis; (c) determining a similarity value

15 above which in said rank ordered list said acceptable number of tumor samples or fewer are false negatives; (d) selecting said similarity value determined in step (c) as a first threshold similarity value; (e) selecting a second similarity value, greater than said first similarity value, as a second threshold similarity value; and (f) determining the similarity between the expression profile of said at least five genes in a breast cancer tumor sample from the breast

20 cancer patient and the expression profile of said respective at least five genes in said pool, to obtain a patient similarity value, wherein if said patient similarity value equals or exceeds said second threshold similarity value, said patient is classified as having a "very good prognosis"; if said patient similarity value equals or exceeds said first threshold similarity value, but is less than said second threshold similarity value, said patient is classified as

25 having an "intermediate prognosis"; and if said patient similarity value is less than said first threshold similarity value, said patient is classified as having a "poor prognosis." Another specific embodiment of this method comprises determining the estrogen receptor (ER) status of said patient, wherein if said patient is ER positive and lymph node negative, said therapeutic regimen assigned to said patient additionally comprises adjuvant hormonal

30 therapy.

The invention provides a method of predicting the responsiveness of a breast cancer patient to chemotherapy based on the patient's cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from the patient, the

patient's estrogen receptor level ("ER level"), and the patient's estrogen receptor level relative to the patient's age ("ER/AGE"). In the method of the invention, a patient is classified into an appropriate chemotherapy responsiveness group as compared to patients in the general population of breast cancer patients. The inventors have discovered that a patient

5 exhibits enhanced response to chemotherapy, e.g., to treatment by a combination of cyclophosphamide, methotrexate, and 5-fluorouracil (the "CMF combination"), as compared to patients in the general population of breast cancer patients if the patient's cellular constituent profile indicates poor prognosis, and the patient's ER level is ER⁺, and the patient's estrogen receptor level relative to the patient's age ER/AGE is low, whereas the

10 patient exhibits reduced response to chemotherapy as compared to patients in the general population if (i) the patient's ER level is ER⁻, or (ii) the patient's cellular constituent profile indicates poor prognosis, and the patient's ER level is ER⁺, and the patient's estrogen receptor level relative to the patient's age ER/AGE is high. Here, ER⁺ designates a high ER level and ER⁻ designates a low ER level. In one embodiment, the ER levels are measured by

15 immunohistochemical staining with tissue samples, and those patients whose tissue samples show greater than 10% of nuclei showing staining is deemed ER⁺ and less than 10% of nuclei showing staining is deemed ER⁻. In another embodiment, ER⁺ and ER⁻ patients are separated by the transcript or mRNA level of a gene encoding the estrogen receptor measured using a microarray, with ER⁺ patients having $\log_{10}(\text{ratio}) > -0.65$, and ER⁻ patients having

20 $\log_{10}(\text{ratio}) =$ or < -0.65 , where the ratio is the ratio of the estrogen receptor mRNA level of the patient sample and a control sample, e.g., a pool of cellular constituents from a plurality of different breast tumor samples. Preferably, the plurality of samples comprises at least 50, 100, 200, or 300 different samples. In addition, the inventors have also found that if the patient's cellular constituent profile indicates a good prognosis, then the patient does not need

25 chemotherapy. The method of the invention is particularly useful for predicting the responsiveness of a breast cancer patient under the age of 55 to chemotherapy, e.g., treatment by the CMF combination.

In one embodiment, the patient's cellular constituent profile is evaluated to determine whether the profile indicates good prognosis or poor prognosis. In a preferred embodiment,

30 the patient's prognosis is evaluated by comparing the cellular constituent profile to a predetermined cellular constituent template profile corresponding to a certain prognosis level, e.g., a good prognosis template comprising measurements of the plurality of cellular constituents which are representative of levels of the cellular constituents in a plurality of

good outcome patients or a poor prognosis template comprising measurements of the plurality of cellular constituents which are representative of levels of the cellular constituents in a plurality of poor outcome patients. In a preferred embodiment, the good prognosis template comprises average levels of the respective cellular constituents in a plurality of good outcome patients. In a preferred embodiment, the poor prognosis template comprises average levels of the respective cellular constituents in a plurality of poor outcome patients. In one embodiment, the average level of each cellular constituent in the good prognosis or poor prognosis template is a simple average. In another embodiment, the average level of each cellular constituent in the good prognosis or poor prognosis template is an error-weighted average. Herein a good outcome patient is a patient who has non-occurrence of metastases within a period of time after initial diagnosis, e.g., a period of 1, 2, 3, 4, 5 or 10 years, and a poor outcome patient is a patient who has occurrence of metastases within a period of time after initial diagnosis, e.g., a period of 1, 2, 3, 4, 5 or 10 years. In a preferred embodiment, both periods are 10 years. Table 8 shows exemplary good (column C_1) and poor (column C_2) template profiles for the 70-gene marker set.

The degree of similarity of the patient's cellular constituent profile to a template representing good or poor prognosis can be used to indicate whether the patient has good or poor prognosis. In a preferred embodiment, a patient is classified as having a good prognosis profile if the patient's cellular constituent profile has a high similarity to a good prognosis template and/or has a low similarity to a poor prognosis template. In another embodiment, a patient is classified as having a poor prognosis profile if the patient's cellular constituent profile has a low similarity to a good prognosis template and/or has a high similarity to a poor prognosis template. In embodiments for predicting the responsiveness of a breast cancer patient under the age of 55, the patients in the good and poor outcome patient populations used to generate the templates are preferably also under the age of 55 at the time of diagnosis of breast cancer.

The degree of similarity between a patient's cellular constituent profile and a template profile can be determined using any method known in the art. In one embodiment, the similarity is represented by a correlation coefficient between the patient's profile and the template. In one embodiment, a correlation coefficient above a correlation threshold indicates high similarity, whereas a correlation coefficient below the threshold indicates low similarity. In preferred embodiments, the correlation threshold is set as 0.3, 0.4, 0.5 or 0.6.

In another embodiment, similarity between a patient's profile and a template is represented by a distance between the patient's profile and the template. In one embodiment, a distance below a given value indicates high similarity, whereas a distance equal to or greater than the given value indicates low similarity.

- 5 As an illustration, in one embodiment, a template for a good prognosis is defined as \bar{c}_1 (e.g., a profile consisting of the values listed in the good prognosis C_1 column of Table 8) and/or a template for poor prognosis is defined as \bar{c}_2 (e.g., a profile consisting of the values listed in the good prognosis C_2 column of Table 8). Either one or both of the two classifier parameters (P_1 and P_2) can then be used to measure degrees of similarities between a
- 10 patient's profile and the templates: P_1 measures the similarity between the patient's profile \bar{y} and the good prognosis template \bar{c}_1 , and P_2 measures the similarity between \bar{y} and the poor prognosis template \bar{c}_2 . In embodiments which employ correlation coefficients, the correlation coefficients P_i can be calculated as:

$$P_i = (\bar{c}_i \cdot \bar{y}) / (\|\bar{c}_i\| \cdot \|\bar{y}\|) \quad (4)$$

- 15 where $i = 1$ and 2.

Thus, in one embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than a selected correlation threshold or if P_2 is equal to or less than a selected correlation threshold. In another embodiment, \bar{y} is classified as a poor prognosis profile if P_1 is less than a selected correlation threshold or if P_2 is above a selected correlation threshold. In still another

20 embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than a first selected correlation threshold and \bar{y} is classified as a poor prognosis profile if P_2 is greater than a second selected correlation threshold.

Other methods known in the art, e.g., the methods described in Section 5.2.4, *infra*, can also be used for classifying an expression profile according to prognosis.

- 25 In a preferred embodiment, the cellular constituent profile is an expression profile comprising measurements of a plurality of transcripts (measured e.g., by measuring mRNAs or cDNA) in a sample derived from a patient. In this embodiment, the good prognosis template can be a good prognosis expression template comprising measurements of the

plurality of transcripts which are representative of expression levels of the transcripts in a plurality of good outcome patients, and the poor prognosis template can be a poor prognosis expression template comprising measurements of the plurality of transcripts which are representative of expression levels of the transcripts in a plurality of poor outcome patients.

- 5 In a preferred embodiment, measurement of each transcript in the good or poor prognosis expression template is an average of expression levels of the transcript in the plurality of good or poor outcome patients, respectively. In one embodiment, each measurement is a mean subtracted log(intensity) (see, PCT publication WO 2004/065545, published August 5, 2004, which is incorporated herein by reference in its entirety). In one embodiment, the
- 10 plurality of transcripts consists of transcripts corresponding to at least a portion of the set of genes listed in Table 8. Preferably, the plurality of transcripts consists of transcripts corresponding to at least 50%, 60%, 70%, 80%, 85%, 90%, 95% or 98% of the set of genes listed in Table 8, or at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, or 65 of the genes listed in Table 8. In a preferred embodiment, the plurality of transcripts consists of
- 15 transcripts corresponding to all the genes listed in Table 8.

- In another embodiment, the expression profile is a differential expression profile comprising differential measurements of said plurality of transcripts in a sample derived from said patient versus measurements of said plurality of transcripts in a control sample. The differential measurements can be $xdev$, $\log(\text{ratio})$, error-weighted $\log(\text{ratio})$, or a mean
- 20 subtracted $\log(\text{intensity})$ (see, e.g., PCT publication WO 00/39339, published on July 6, 2000; PCT publication WO 2004/065545, published August 5, 2004, each of which is incorporated herein by reference in its entirety).

- In another embodiment, the cellular constituent profile comprises measurements of a plurality of protein species in a sample derived from the patient. In this embodiment, a good
- 25 prognosis template comprises measurements of the plurality of protein species which are representative of levels of the protein species in a plurality of good outcome patients, and the poor prognosis template comprises measurements of the plurality of protein species which are representative of levels of the protein species in the plurality of poor outcome patients.

- The patient's cellular constituent profile can be determined by any method known in
- 30 the art, e.g., as described in Section 5.3.

The methods of the invention preferably use a control or reference sample, which can be any suitable sample against which changes in cellular constituents can be determined. In one embodiment, the control or reference sample is generated by pooling together the plurality of cellular constituents, e.g., a plurality of transcripts or cDNAs, or a plurality of protein species, from a plurality of breast cancer patients. Alternatively, the control or reference sample can be generated by pooling together purified or synthesized cellular constituents, e.g., a plurality of purified or synthesized transcripts or cDNAs, a plurality of purified or synthesized protein species. In one embodiment, synthetic RNAs for each transcripts or cDNAs are pooled to form the control or reference sample. Preferably, the abundances of synthetic RNAs are approximately the abundances of the corresponding transcripts in a real tumor pool. The differential expression of marker genes for each individual patient sample is measured against this control sample. In one embodiment, 60-mer oligonucleotides corresponding to the probe sequences on a microarray used to assay the expression levels of the diagnostic/prognostic transcripts are synthesized and cloned into pBluescript SK- vector (Statagene, La Jolla, CA), adjacent to the T7 promotor sequence. Individual clones are isolated, and the sequences of their inserts are verified by DNA sequencing. To generate synthetic RNAs, clones are linearized with *EcoRI* and a T7 in vitro transcription (IVT) reaction is performed by MegaScript kit (Ambion, Austin, TX), followed by DNase treatment of the product. Synthetic RNAs are purified on RNeasy columns (Qiagen, Valencia, CA). These synthetic RNAs are transcribed, amplified, labeled, and mixed together to make the reference pool. The abundance of those synthetic RNAs are chosen to approximate the abundances of the transcripts of the corresponding marker genes in the real tumor pool.

In one embodiment, the ER level is determined by measuring an expression level of a gene encoding the estrogen receptor in the patient relative to that in a control sample. In a preferred embodiment, the gene encoding the estrogen receptor is the estrogen receptor α gene. In another embodiment, the ER level is measured by an oligonucleotide probe that detects the transcript of the gene having accession number NM_000125 (SEQ ID NO: 425). The control sample is a pool of breast cancer cells from different patients. In this embodiment, the ER level is classified as ER⁺ if the $\log(\text{ratio}) > -0.65$, and is classified as ER⁻ otherwise (see, e.g., van't Veer et al., 2002, Nature 415:530). In another embodiment, the ER level is determined by immunohistochemical staining measurement of the level of the estrogen receptors in a tissue sample of the patient. In one embodiment, the ER level is

classified as ER⁺ if immunohistochemical staining shows greater than 10% of nuclei showing staining and the ER level is classified as ER⁻, if less than 10% nuclei of showing staining.

The patient's estrogen receptor level relative to the patient's age can be measured using an appropriate metric designated as ER/AGE. If the individual's ER level is high
5 relative to the individual's age, e.g., if the expression level of an estrogen receptor gene of the individual is higher than an age-dependent threshold value, the individual is classified as ER/AGE high. In one embodiment, the log(ratio) of ER expression of an individual of a particular age is compared with the threshold value for the individual's age, and an ER expression level equal to or greater than the threshold value is deemed high. In one
10 embodiment, the patient's estrogen receptor level as measured by log₁₀(ratio) is compared to a parameter defined as $c \cdot (AGE - d)$, where c is a coefficient, AGE is the age of said patient, and d is an age threshold. The patient is classified as having high ER/AGE if the patient's ER level is greater than $c \cdot (AGE - d)$, and the patient is classified as having low ER/AGE if the patient's ER level is equal to or less than $c \cdot (AGE - d)$. The parameters c and d depend on
15 the ER level and AGE used. They can be determined by fitting patients' ER level-age distribution to a bimodal distribution of two subgroups each having a different ER level-age dependence. In a preferred embodiment, $c = 0.1$ and $d = 42.5$.

Chemotherapy can be performed using any one or a combination of the anti-cancer drugs known in the art, including but are not limited to any topoisomerase inhibitor, DNA
20 binding agent, anti-metabolite, ionizing radiation, or a combination of two or more of such known DNA damaging agents.

A topoisomerase inhibitor that can be used in conjunction with the invention can be a topoisomerase I (Topo I) inhibitor, a topoisomerase II (Topo II) inhibitor, or a dual topoisomerase I and II inhibitor. A topo I inhibitor can be from any of the following classes
25 of compounds: camptothecin analogue (e.g., karenitecin, aminocamptothecin, lurtotecan, topotecan, irinotecan, BAY 56-3722, rubitecan, GI14721, exatecan mesylate), rebeccamycin analogue, PNU 166148, rebeccamycin, TAS-103, camptothecin (e.g., camptothecin polyglutamate, camptothecin sodium), intoplicine, ecteinascidin 743, J-107088, pibenzimol. Examples of preferred topo I inhibitors include but are not limited to camptothecin, topotecan
30 (hycaptamine), irinotecan (irinotecan hydrochloride), belotecan, or an analogue or derivative thereof.

A topo II inhibitor that can be used in conjunction with the invention can be from any of the following classes of compounds: anthracycline antibiotics (e.g., carubicin, pirarubicin, daunorubicin citrate liposomal, daunomycin, 4-iodo-4-doxydoxorubicin, doxorubicin, n,n-dibenzyl daunomycin, morpholinodoxorubicin, aclacinomycin antibiotics, duborimycin, 5 menogaril, nogalamycin, zorubicin, epirubicin, marcellomycin, detorubicin, annamycin, 7-cyanoquinocarcinol, deoxydoxorubicin, idarubicin, GPX-100, MEN-10755, valrubicin, KRN5500), epipodophyllotoxin compound (e.g., podophyllin, teniposide, etoposide, GL331, 2-ethylhydrazide), anthraquinone compound (e.g., ametantrone, bisantrene, mitoxantrone, anthraquinone), ciprofloxacin, acridine carboxamide, amonafide, anthrapyrazole antibiotics 10 (e.g., teloxantrone, sedoxantrone trihydrochloride, piroxantrone, anthrapyrazole, losoxantrone), TAS-103, fostriecin, razoxane, XK469R, XK469, chloroquinoxaline sulfonamide, merbarone, intoplicine, elsamitucin, CI-921, pyrazoloacridine, elliptinium, amsacrine. Examples of preferred topo II inhibitors include but are not limited to doxorubicin (Adriamycin), etoposide phosphate (etopofos), teniposide, sobuzoxane, or an 15 analogue or derivative thereof.

DNA binding agents that can be used in conjunction with the invention include but are not limited to DNA groove binding agent, e.g., DNA minor groove binding agent; DNA crosslinking agent; intercalating agent; and DNA adduct forming agent. A DNA minor groove binding agent can be an anthracycline antibiotic, mitomycin antibiotic (e.g., 20 porfiromycin, KW-2149, mitomycin B, mitomycin A, mitomycin C), chromomycin A3, carzelesin, actinomycin antibiotic (e.g., cactinomycin, dactinomycin, actinomycin F1), brostallicin, echinomycin, bizelesin, duocarmycin antibiotic (e.g., KW 2189), adozelesin, olivomycin antibiotic, plicamycin, zinostatin, distamycin, MS-247, ecteinascidin 743, amsacrine, anthramycin, and pibenzimol, or an analogue or derivative thereof.

25 DNA crosslinking agents include but are not limited to antineoplastic alkylating agent, methoxsalen, mitomycin antibiotic, psoralen. An antineoplastic alkylating agent can be a nitrosourea compound (e.g., cysteamustine, tauromustine, semustine, PCNU, streptozocin, SarCNU, CGP-6809, carmustine, fotemustine, methyl nitrosourea, nimustine, ranimustine, ethyl nitrosourea, lomustine, chlorozotocin), mustard agent (e.g., nitrogen 30 mustard compound, such as spiromustine, trofosfamide, chlorambucil, estramustine, 2,2,2-trichlorotriethylamine, prednimustine, novembichin, phenamet, glufosfamide, peptichemio, ifosfamide, defosfamide, nitrogen mustard, phenesterin, mannomustine, cyclophosphamide,

melphalan, perfosfamide, mechlorethamine oxide hydrochloride, uracil mustard, bestrabucil, DHEA mustard, tallimustine, mafosfamide, aniline mustard, chlornaphazine; sulfur mustard compound, such as bischloroethylsulfide; mustard prodrug, such as TLK286 and ZD2767), ethylenimine compound (e.g., mitomycin antibiotic, ethylenimine, uredepa, thiotepa, diaziquone, hexamethylene bisacetamide, pentamethylmelamine, altretamine, carzinophilin, triaziquone, meturedopa, benzodepa, carboquone), alkylsulfonate compound (e.g., dimethylbusulfan, Yoshi-864, improsulfan, piposulfan, treosulfan, busulfan, hepsulfam), epoxide compound (e.g., anaxirone, mitolactol, dianhydrogalactitol, teroxirone), miscellaneous alkylating agent (e.g., ipomeanol, carzelesin, methylene dimethane sulfonate, mitobronitol, bizelesin, adozelesin, piperazinedione, VNP40101M, asaley, 6-hydroxymethylacetylfulvene, EO9, etoglucid, ecteinascidin 743, pipobroman), platinum compound (e.g., ZD0473, liposomal-cisplatin analogue, satraplatin, BBR 3464, spiroplatin, ormaplatin, cisplatin, oxaliplatin, carboplatin, lobaplatin, zeniplatin, iproplatin), triazene compound (e.g., imidazole mustard, CB10-277, mitozolomide, temozolomide, procarbazine, dacarbazine), picoline compound (e.g., penclomedine), or an analogue or derivative thereof. Examples of preferred alkylating agents include but are not limited to cisplatin, dibromodulcitol, fotemustine, ifosfamide (ifosfamid), ranimustine (ranomustine), nedaplatin (latoplatin), bendamustine (bendamustine hydrochloride), eptaplatin, temozolomide (methazolastone), carboplatin, altretamine (hexamethylmelamine), prednimustine, oxaliplatin (oxalaplatinum), carmustine, thiotepa, leusulfon (busulfan), lobaplatin, cyclophosphamide, bisulfan, melphalan, and chlorambucil, or analogues or derivatives thereof.

Intercalating agents can be an anthraquinone compound, bleomycin antibiotic, rebeccamycin analogue, acridine, acridine carboxamide, amonafide, rebeccamycin, anthracycline antibiotic, echinomycin, psoralen, LU 79553, BW A773U, crisnatol mesylate, benzo(a)pyrene-7,8-diol-9,10-epoxide, acodazole, elliptinium, pixantrone, or an analogue or derivative thereof.

DNA adduct forming agents include but are not limited to enediyne antitumor antibiotic (e.g., dynemicin A, esperamicin A1, zinostatin, dynemicin, calicheamicin gamma 1I), platinum compound, carmustine, tamoxifen (e.g., 4-hydroxy-tamoxifen), psoralen, pyrazine diazohydroxide, benzo(a)pyrene-7,8-diol-9,10-epoxide, or an analogue or derivative thereof.

Anti-metabolites include but are not limited to cytosine, arabinoside, floxuridine, fluorouracil, mercaptopurine, Gemcitabine, and methotrexate (MTX).

In a specific embodiment, a KSP inhibitor (1S)-1-{[(2S)-4-(2,5-difluorophenyl)-2-phenyl-2,5-dihydro-1H-pyrrol-1-yl]carbonyl}-2-methylpropylamine (see, PCT application
5 PCT/US03/18482, filed June 12, 2003, which is incorporated herein by reference in its entirety), is used alone or in combination with other anti-cancer drugs for chemotherapy.

In another specific embodiment, the chemotherapy is carried out using the CMF combination consisting of cyclophosphamide, methotrexate, and 5-fluorouracil.

5.2.3. IMPROVING SENSITIVITY TO EXPRESSION LEVEL DIFFERENCES

10 In using the markers disclosed herein, and, indeed, using any sets of markers, e.g., to compare profiles or to differentiate an individual having one phenotype from another individual having a second phenotype, one can compare the profile comprising absolute expression levels of the markers in a sample to a template; for example, a template comprising the average levels of expression of the markers in a plurality of individuals. To
15 increase the sensitivity of the comparison, however, the expression level values are preferably transformed in a number of ways. Also, to differentiate an individual having one phenotype from another individual having a second phenotype using any sets of markers, one can compare the absolute expression of each of the markers in a sample to a control; for example, the control can be the average level of expression of each of the markers, respectively, in a
20 pool of individuals.

For example, the expression level of each of the markers can be normalized by the average expression level of all markers the expression level of which is determined, or by the average expression level of a set of control genes. Thus, in one embodiment, the markers are represented by probes on a microarray, and the expression level of each of the markers is
25 normalized by the mean or median expression level across all of the genes represented on the microarray, including any non-marker genes. In a specific embodiment, the normalization is carried out by dividing the median or mean level of expression of all of the genes on the microarray. In another embodiment, the expression levels of the markers is normalized by the mean or median level of expression of a set of control markers. In a specific
30 embodiment, the control markers comprise a set of housekeeping genes. In another specific

embodiment, the normalization is accomplished by dividing by the median or mean expression level of the control genes.

The sensitivity of a marker-based assay will also be increased if the expression levels of individual markers are compared to the expression of the same markers in a pool of samples. Preferably, the comparison is to the mean or median expression level of each the marker genes in the pool of samples. Such a comparison may be accomplished, for example, by dividing by the mean or median expression level of the pool for each of the markers from the expression level each of the markers in the sample. This has the effect of accentuating the relative differences in expression between markers in the sample and markers in the pool as a whole, making comparisons more sensitive and more likely to produce meaningful results that the use of absolute expression levels alone. The expression level data may be transformed in any convenient way; preferably, the expression level data for all is log transformed before means or medians are taken.

In performing comparisons to a pool, two approaches may be used. First, the expression levels of the markers in the sample may be compared to the expression level of those markers in the pool, where nucleic acid derived from the sample and nucleic acid derived from the pool are hybridized during the course of a single experiment. Such an approach requires that new pool nucleic acid be generated for each comparison or limited numbers of comparisons, and is therefore limited by the amount of nucleic acid available. Alternatively, and preferably, the expression levels in a pool, whether normalized and/or transformed or not, are stored on a computer, or on computer-readable media, to be used in comparisons to the individual expression level data from the sample (i.e., single-channel data).

The current invention also provides the following method of classifying a first cell or organism as having one of at least two different phenotypes, where the different phenotypes comprise a first phenotype and a second phenotype. The level of expression of each of a plurality of markers in a first sample from the first cell or organism is compared to the level of expression of each of said markers, respectively, in a pooled sample from a plurality of cells or organisms, the plurality of cells or organisms comprising different cells or organisms exhibiting said at least two different phenotypes, respectively, to produce a first compared value. The first compared value is then compared to a second compared value, wherein said second compared value is the product of a method comprising comparing the level of

expression of each of said markers in a sample from a cell or organism characterized as having said first phenotype to the level of expression of each of said markers, respectively, in the pooled sample. The first compared value is then compared to a third compared value, wherein said third compared value is the product of a method comprising comparing the level of expression of each of the markers in a sample from a cell or organism characterized as having the second phenotype to the level of expression of each of the markers, respectively, in the pooled sample. In specific embodiments, the marker can be a gene, a protein encoded by the gene, etc. Optionally, the first compared value can be compared to additional compared values, respectively, where each additional compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having a phenotype different from said first and second phenotypes but included among the at least two different phenotypes, to the level of expression of each of said genes, respectively, in said pooled sample. Finally, a determination is made as to which of said second, third, and, if present, one or more additional compared values, said first compared value is most similar, wherein the first cell or organism is determined to have the phenotype of the cell or organism used to produce said compared value most similar to said first compared value.

The sensitivity of a marker-based assay will also be increased if the expression levels of individual markers are compared to the expression of the same markers in a control sample, e.g., a sample comprises a pool of samples, to generate a differential expression profile. Such a comparison may be accomplished, for example, by determining a ratio between expression level of each marker in the sample and the expression level of the corresponding marker in the control sample. This has the effect of accentuating the relative differences in expression between markers in the sample and markers in the control as a whole, making subsequent comparisons to a template more sensitive and more likely to produce meaningful results than the use of absolute expression levels alone. The comparison may be performed in any convenient way, e.g., by taking difference, ratio, or log(ratio).

In performing comparisons to a control sample, two approaches may be used. First, the expression levels of the markers in the sample may be compared to the expression level of those markers in the control sample, e.g., where nucleic acid derived from the sample and nucleic acid derived from the control are hybridized during the course of a single experiment. Such an approach requires that new control sample of nucleic acid be generated for each

comparison or limited numbers of comparisons, and is therefore limited by the amount of nucleic acid available. Alternatively, the expression levels in a control sample, whether normalized and/or transformed or not, are stored on a computer, or on computer-readable media, to be used in comparisons to the individual expression level data from the sample (i.e., single-channel data).

The methods of the invention preferably use a control or reference sample, which can be any suitable sample against which changes in cellular constituents can be determined. In one embodiment, the control or reference sample is generated by pooling together the plurality of cellular constituents, e.g., a plurality of transcripts or cDNAs, or a plurality of protein species, from a plurality of breast cancer patients. Alternatively, the control or reference sample can be generated by pooling together purified or synthesized cellular constituents, e.g., a plurality of purified or synthesized transcripts or cDNAs, a plurality of purified or synthesized protein species. In one embodiment, synthetic RNAs for each transcripts or cDNAs are pooled to form the control or reference sample. Preferably, the abundances of synthetic RNAs are approximately the abundances of the corresponding transcripts in a real tumor pool. The differential expression of marker genes for each individual patient sample is measured against this control sample. In one embodiment, 60-mer oligonucleotides corresponding to the probe sequences on a microarray used to assay the expression levels of the diagnostic/prognostic transcripts are synthesized and cloned into pBluescript SK- vector (Statagene, La Jolla, CA), adjacent to the T7 promotor sequence. Individual clones are isolated, and the sequences of their inserts are verified by DNA sequencing. To generate synthetic RNAs, clones are linearized with *EcoRI* and a T7 in vitro transcription (IVT) reaction is performed by MegaScript kit (Ambion, Austin, TX), followed by DNase treatment of the product. Synthetic RNAs are purified on RNeasy columns (Qiagen, Valencia, CA). These synthetic RNAs are transcribed, amplified, labeled, and mixed together to make the reference pool. The abundance of those synthetic RNAs are chosen to approximate the abundances of the transcripts of the corresponding marker genes in the real tumor pool.

The current invention provides the following method of classifying a first cell or organism as having one of at least two different phenotypes, where the different phenotypes comprise a first phenotype and a second phenotype. The level of expression of each of a plurality of markers in a first sample from the first cell or organism is compared to the level

of expression of each of said markers, respectively, in a pooled sample from a plurality of cells or organisms, the plurality of cells or organisms comprising different cells or organisms exhibiting said at least two different phenotypes, respectively, to produce a first compared value so that a first differential profile comprising a plurality of first compared values for said plurality of markers is generated. The first differential profile is then compared to a second differential profile comprising a plurality of second compared values, wherein each said second compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having said first phenotype to the level of expression of each of said markers, respectively, in the pooled sample. The first differential profile is then compared to a third differential profile comprising a plurality of third compared values, wherein each said third compared value is the product of a method comprising comparing the level of expression of each of the markers in a sample from a cell or organism characterized as having the second phenotype to the level of expression of each of the markers, respectively, in the pooled sample. In specific embodiments, each marker can be a gene, a protein encoded by the gene, etc. Optionally, the first differential profile can be compared to additional expression profiles each of which comprising additional compared values, respectively, where each additional compared value is the product of a method comprising comparing the level of expression of each of said markers in a sample from a cell or organism characterized as having a phenotype different from said first and second phenotypes but included among the at least two different phenotypes, to the level of expression of each of said genes, respectively, in said pooled sample. Finally, a determination is made as to which of said second, third, and, if present, one or more additional differential profiles, said first differential profile is most similar, wherein the first cell or organism is determined to have the phenotype of the cell or organism used to produce said differential profile most similar to said first differential profile.

In a specific embodiment of this method, the compared values are each ratios of the levels of expression of each of said genes. In another specific embodiment, each of the levels of expression of each of the genes in the pooled sample is normalized prior to any of the comparing steps. In a more specific embodiment, the normalization of the levels of expression is carried out by dividing by the median or mean level of the expression of each of the genes or dividing by the mean or median level of expression of one or more housekeeping genes in the pooled sample from said cell or organism. In another specific embodiment, the normalized levels of expression are subjected to a log transform, and the comparing steps

comprise subtracting the log transform from the log of the levels of expression of each of the genes in the sample. In another specific embodiment, the two or more different phenotypes are different stages of a disease or disorder. In still another specific embodiment, the two or more different phenotypes are different prognoses of a disease or disorder. In yet another specific embodiment, the levels of expression of each of the genes, respectively, in the pooled sample or said levels of expression of each of said genes in a sample from the cell or organism characterized as having the first phenotype, second phenotype, or said phenotype different from said first and second phenotypes, respectively, are stored on a computer or on a computer-readable medium.

10 In another specific embodiment, the two phenotypes are good prognosis and poor prognosis. In a more specific embodiment, the two phenotypes are good prognosis and poor prognosis for an individual that is identified as having ER⁻, *BRCA1* status, ER⁻, sporadic status, ER⁺, ER/AGE high status, ER⁺, ER/AGE low, LN⁺ status, or ER⁺, ER/AGE low, LN⁺ status.

15 In another specific embodiment, the comparison is made between the expression profile of the genes in the sample and the expression profile of the same genes in a pool representing only one of two or more phenotypes. In the context of prognosis-correlated genes, for example, one can compare the expression levels of prognosis-related genes in a sample to the average levels of the expression of the same genes in a plurality of "good prognosis" samples (as opposed to a plurality of samples that include samples from patients having poor prognoses and good prognoses). Thus, in this method, a sample is classified as having a good prognosis if the expression profile of prognosis-correlated genes exceeds a chosen coefficient of correlation to the average "good prognosis" expression profile (*e.g.*, the profile comprising average levels of expression of prognosis-correlated genes in samples from a plurality of patients having a "good prognosis"). Patients whose expression profiles correlate more poorly with the "good prognosis" expression profile (*e.g.*, whose correlation coefficient fails to exceed the chosen coefficient) are classified as having a poor prognosis.

Where individuals are classified on the basis of phenotypic, genotypic, or clinical characteristics into patient subsets, the pool of samples may be a pool of samples for the phenotype that includes samples representing each of the patient subsets. Alternatively, the pool of samples may be a pool of samples for the phenotype representing only the specific patient subset. For example, where an individual is classified as ER⁺, sporadic, the pool of

samples to which the individual's sample is compared may be a pool of samples from ER+, sporadic individuals having a good prognosis only, or may be a pool of samples of individuals having a good prognosis, without regard to ER status or mutation type.

The method can be applied to a plurality of patient subsets. For example, in a specific embodiment, the phenotype is good prognosis, and the individual is classified into one of the following patient subsets: ER⁻, *BRCA1* status, ER⁻, sporadic status, ER+, ER/AGE high status, ER+, ER/AGE low, LN+ status, or ER+, ER/AGE low, LN+ status. A set of markers informative for prognosis for the patient subset into which the individual is classified is then used to determine the likely prognosis for the individual. A sample is classified as coming from an individual having a good prognosis if the expression profile of prognosis-correlated genes for the particular subset into which the individual is classified exceeds a chosen coefficient of correlation to the average "good prognosis" expression profile (e.g., the levels of expression of prognosis-correlated genes in a plurality of samples from patients within the subclass having a "good prognosis"). Patients whose expression levels correlate more poorly with the "good prognosis" expression profile (e.g., whose correlation coefficient fails to exceed the chosen coefficient) are classified as having a poor prognosis.

Of course, single-channel data may also be used without specific comparison to a mathematical sample pool. For example, a sample may be classified as having a first or a second phenotype, wherein the first and second phenotypes are related, by calculating the similarity between the expression profile of at least 5 markers in the sample, where the markers are correlated with the first or second phenotype, to a first phenotype template and a second phenotype template each comprising the expression levels of the same markers, by (a) labeling nucleic acids derived from a sample with a fluorophore to obtain a pool of fluorophore-labeled nucleic acids; (b) contacting said fluorophore-labeled nucleic acid with a microarray under conditions such that hybridization can occur, detecting at each of a plurality of discrete loci on the microarray a fluorescent emission signal from said fluorophore-labeled nucleic acid that is bound to said microarray under said conditions; and (c) determining the similarity of marker gene expression in the individual sample to the first and second templates, wherein if said expression is more similar to the first template, the sample is classified as having the first phenotype, and if said expression is more similar to the second template, the sample is classified as having the second phenotype.

In a specific embodiment of the above method, the first phenotype is a good prognosis of breast cancer, the sample is a sample from an individual that has been classified into a patient subset, and the first and second templates are templates for the phenotype for the particular patient subset. In a more specific embodiment, for example, the first phenotype is a good prognosis, the second phenotype is a poor prognosis, the patient is classified into an ER⁻, sporadic patient subset, an ER⁻, *BRCA1* subset, an ER⁺, ER/AGE high subset, an ER⁺, ER/AGE low, LN⁺ subset, or an ER⁺, ER/AGE low, LN⁺ subset, and said first and second templates are templates derived from the expression of the marker genes in individuals having a good prognosis and a poor prognosis, respectively, wherein said individuals are all of the patient subset into which said patient is classified.

5.2.4. METHODS FOR CLASSIFICATION OF EXPRESSION PROFILES

In preferred embodiments, the methods of the invention use a classifier for predicting prognosis in a patient. The classifier can be based on any appropriate pattern recognition method that receives an input comprising a marker profile and provides an output comprising data indicating which patient subset the patient belongs. The classifier can be trained with training data from a training population of breast cancer patients. Typically, the training data comprise for each of the patients in the training population a training marker profile comprising measurements of respective gene products of a plurality of genes in a suitable sample taken from the patient and prognosis information.

In preferred embodiments, the classifier can be based on a classification (pattern recognition) method described below, e.g., profile similarity (Section 5.2.4.1., *infra*); artificial neural network (Section 5.2.4.2., *infra*); support vector machine (SVM, Section 5.2.4.3., *infra*); logic regression (Section 5.2.4.4., *infra*), linear or quadratic discriminant analysis (Section 5.2.4.5., *infra*), decision trees (Section 5.2.4.6., *infra*), clustering (Section 5.2.4.7., *infra*), principal component analysis (Section 5.2.4.8., *infra*), nearest neighbor classifier analysis (Section 5.2.4.9., *infra*). Such classifiers can be trained with the training population using methods described in the relevant sections, *infra*.

The marker profile can be obtained by measuring the plurality of gene products in a cell sample from the patient using a method known in the art, e.g., a method described in Section 5.3., *infra*.

Various known statistical pattern recognition methods can be used in conjunction with the present invention. A classifier based on any of such methods can be constructed using the marker profiles and prognosis data of training patients. Such a classifier can then be used to evaluate the prognosis status of a patient based on the patient's marker profile. The methods
 5 can also be used to identify markers that discriminate between different prognosis status using a marker profile and prognosis data of training patients.

5.2.4.1. PROFILE MATCHING

A patient's prognosis can be evaluated by comparing a marker profile obtained in a suitable sample from the patient with a marker profile that is representative of a particular
 10 prognosis. Such a marker profile is also termed a "template profile" or a "template." The degree of similarity to such a template profile provides an evaluation of the patient's prognosis. If the degree of similarity of the patient marker profile and a template profile is above a predetermined threshold, the patient is assigned the prognosis represented by the template. For example, a patient's prognosis can be evaluated by comparing a marker profile
 15 of the patient to a predetermined template profile corresponding to a given prognosis, e.g., a good prognosis template comprising measurements of the plurality of markers which are representative of levels of the markers in a plurality of patients having a good prognosis.

In one embodiment, the similarity is represented by a correlation coefficient between the patient's profile and the template. In one embodiment, a correlation coefficient above a
 20 correlation threshold indicates a high similarity, whereas a correlation coefficient below the threshold indicates a low similarity.

In a specific embodiment, P_i measures the similarity between the patient's profile \bar{y} and a template profile comprising measurements of marker gene products representative of measurements of marker gene products in patients having a particular prognosis, e.g., the
 25 good prognosis template \bar{z}_1 or the poor prognosis template \bar{z}_2 . Such a coefficient, P_i , can be calculated using the following equation:

$$P_i = (\bar{z}_i \bullet \bar{y}) / (\|\bar{z}_i\| \cdot \|\bar{y}\|)$$

where i designates the i th template. Thus, in one embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than a selected correlation threshold. In another

embodiment, \bar{y} is classified as poor prognosis profile if P_2 is greater than a selected correlation threshold. In preferred embodiments, the correlation threshold is set as 0.3, 0.4, 0.5 or 0.6. In another embodiment, \bar{y} is classified as a good prognosis profile if P_1 is greater than P_2 , whereas \bar{y} is classified as a poor prognosis profile if P_1 is less than P_2 .

5 In another embodiment, the correlation coefficient is a weighted dot product of the patient's profile \bar{y} and a template profile, in which measurements of each different marker is assigned a weight.

In another embodiment, similarity between a patient's profile and a template is represented by a distance between the patient's profile and the template. In one embodiment,
10 a distance below a given value indicates high similarity, whereas a distance equal to or greater than the given value indicates low similarity.

In one embodiment, the Euclidian distance according to the formula

$$D_i = \|\bar{y} - \bar{z}_i\|$$

is used, where D_i measures the distance between the patient's profile \bar{y} and a template
15 profile comprising measurements of marker gene products representative of measurements of marker gene products in patients having a particular prognosis, e.g., the good prognosis template \bar{z}_1 or the poor prognosis template \bar{z}_2 . In other embodiments, the Euclidian distance is squared to place progressively greater weight on cellular constituents that are further apart. In alternative embodiments, the distance measure D_i is the Manhattan distance provide by

20
$$D_i = \sum_n |y(n) - z_i(n)|$$

where $y(n)$ and $z_i(n)$ are respectively measurements of the n th marker gene product in the patient's profile \bar{y} and a template profile.

In another embodiment, the distance is defined as $D_i = 1 - P_i$, where P_i is the correlation coefficient or normalized dot product as described above.

25 In still other embodiments, the distance measure may be the Chebychev distance, the power distance, and percent disagreement, all of which are well known in the art.

5.2.4.2. ARTIFICIAL NEURAL NETWORK

In some embodiments, a neural network is used. A neural network can be constructed for a selected set of molecular markers of the invention. A neural network is a two-stage regression or classification model. A neural network has a layered structure that includes a layer of input units (and the bias) connected by a layer of weights to a layer of output units. For regression, the layer of output units typically includes just one output unit. However, neural networks can handle multiple quantitative responses in a seamless fashion.

In multilayer neural networks, there are input units (input layer), hidden units (hidden layer), and output units (output layer). There is, furthermore, a single bias unit that is connected to each unit other than the input units. Neural networks are described in Duda *et al.*, 2001, *Pattern Classification*, Second Edition, John Wiley & Sons, Inc., New York; and Hastie *et al.*, 2001, *The Elements of Statistical Learning*, Springer-Verlag, New York.

The basic approach to the use of neural networks is to start with an untrained network, present a training pattern, e.g., marker profiles from training patients, to the input layer, and to pass signals through the net and determine the output, e.g., the prognosis in the training patients, at the output layer. These outputs are then compared to the target values; any difference corresponds to an error. This error or criterion function is some scalar function of the weights and is minimized when the network outputs match the desired outputs. Thus, the weights are adjusted to reduce this measure of error. For regression, this error can be sum-of-squared errors. For classification, this error can be either squared error or cross-entropy (deviation). See, e.g., Hastie *et al.*, 2001, *The Elements of Statistical Learning*, Springer-Verlag, New York.

Three commonly used training protocols are stochastic, batch, and on-line. In stochastic training, patterns are chosen randomly from the training set and the network weights are updated for each pattern presentation. Multilayer nonlinear networks trained by gradient descent methods such as stochastic back-propagation perform a maximum-likelihood estimation of the weight values in the model defined by the network topology. In batch training, all patterns are presented to the network before learning takes place. Typically, in batch training, several passes are made through the training data. In online training, each pattern is presented once and only once to the net.

In some embodiments, consideration is given to starting values for weights. If the weights are near zero, then the operative part of the sigmoid commonly used in the hidden layer of a neural network (see, e.g., Hastie *et al.*, 2001, *The Elements of Statistical Learning*, Springer-Verlag, New York) is roughly linear, and hence the neural network collapses into an approximately linear model. In some embodiments, starting values for weights are chosen to be random values near zero. Hence the model starts out nearly linear, and becomes nonlinear as the weights increase. Individual units localize to directions and introduce nonlinearities where needed. Use of exact zero weights leads to zero derivatives and perfect symmetry, and the algorithm never moves. Alternatively, starting with large weights often leads to poor solutions.

Since the scaling of inputs determines the effective scaling of weights in the bottom layer, it can have a large effect on the quality of the final solution. Thus, in some embodiments, at the outset all expression values are standardized to have mean zero and a standard deviation of one. This ensures all inputs are treated equally in the regularization process, and allows one to choose a meaningful range for the random starting weights. With standardization inputs, it is typical to take random uniform weights over the range $[-0.7, +0.7]$.

A recurrent problem in the use of networks having a hidden layer is the optimal number of hidden units to use in the network. The number of inputs and outputs of a network are determined by the problem to be solved. In the present invention, the number of inputs for a given neural network can be the number of molecular markers in the selected set of molecular markers of the invention. The number of output for the neural network will typically be just one. However, in some embodiment more than one output is used so that more than just two states can be defined by the network. If too many hidden units are used in a neural network, the network will have too many degrees of freedom and is trained too long, there is a danger that the network will overfit the data. If there are too few hidden units, the training set cannot be learned. Generally speaking, however, it is better to have too many hidden units than too few. With too few hidden units, the model might not have enough flexibility to capture the nonlinearities in the data; with too many hidden units, the extra weight can be shrunk towards zero if appropriate regularization or pruning, as described below, is used. In typical embodiments, the number of hidden units is somewhere in the

range of 5 to 100, with the number increasing with the number of inputs and number of training cases.

One general approach to determining the number of hidden units to use is to apply a regularization approach. In the regularization approach, a new criterion function is constructed that depends not only on the classical training error, but also on classifier complexity. Specifically, the new criterion function penalizes highly complex models; searching for the minimum in this criterion is to balance error on the training set with error on the training set plus a regularization term, which expresses constraints or desirable properties of solutions:

$$J = J_{pat} + \lambda J_{reg}$$

The parameter λ is adjusted to impose the regularization more or less strongly. In other words, larger values for λ will tend to shrink weights towards zero: typically cross-validation with a validation set is used to estimate λ . This validation set can be obtained by setting aside a random subset of the training population. Other forms of penalty can also be used, for example the weight elimination penalty (see, *e.g.*, Hastie *et al.*, 2001, *The Elements of Statistical Learning*, Springer-Verlag, New York).

Another approach to determine the number of hidden units to use is to eliminate - prune - weights that are least needed. In one approach, the weights with the smallest magnitude are eliminated (set to zero). Such magnitude-based pruning can work, but is nonoptimal; sometimes weights with small magnitudes are important for learning and training data. In some embodiments, rather than using a magnitude-based pruning approach, Wald statistics are computed. The fundamental idea in Wald Statistics is that they can be used to estimate the importance of a hidden unit (weight) in a model. Then, hidden units having the least importance are eliminated (by setting their input and output weights to zero). Two algorithms in this regard are the *Optimal Brain Damage* (OBD) and the *Optimal Brain Surgeon* (OBS) algorithms that use second-order approximation to predict how the training error depends upon a weight, and eliminate the weight that leads to the smallest increase in training error.

Optimal Brain Damage and Optimal Brain Surgeon share the same basic approach of training a network to local minimum error at weight w , and then pruning a weight that leads

to the smallest increase in the training error. The predicted functional increase in the error for a change in full weight vector $\delta \mathbf{w}$ is:

$$\delta J = \left(\frac{\partial J}{\partial \mathbf{w}} \right)' \cdot \delta \mathbf{w} + \frac{1}{2} \delta \mathbf{w}' \cdot \frac{\partial^2 J}{\partial \mathbf{w}^2} \cdot \delta \mathbf{w} + O(\|\delta \mathbf{w}\|^3)$$

where $\frac{\partial^2 J}{\partial \mathbf{w}^2}$ is the Hessian matrix. The first term vanishes because we are at a local

- 5 minimum in error; third and higher order terms are ignored. The general solution for minimizing this function given the constraint of deleting one weight is:

$$\delta w = -\frac{w_q}{[\mathbf{H}^{-1}]_{qq}} \mathbf{H}^{-1} \cdot \mathbf{u}_q \text{ and } L_q = \frac{1}{2} - \frac{w_q^2}{[\mathbf{H}^{-1}]_{qq}}$$

- Here, \mathbf{u}_q is the unit vector along the q th direction in weight space and L_q is approximation to the saliency of the weight q - the increase in training error if weight q is pruned and the other weights updated $\delta \mathbf{w}$. These equations require the inverse of \mathbf{H} . One method to calculate this
- 10 inverse matrix is to start with a small value, $H_0^{-1} = \alpha^{-1} \mathbf{I}$, where α is a small parameter - effectively a weight constant. Next the matrix is updated with each pattern according to

$$\mathbf{H}_{m+1}^{-1} = \mathbf{H}_m^{-1} - \frac{\mathbf{H}_m^{-1} \mathbf{X}_{m+1} \mathbf{X}_{m+1}^T \mathbf{H}_m^{-1}}{\frac{n}{a_m} + \mathbf{X}_{m+1}^T \mathbf{H}_m^{-1} \mathbf{X}_{m+1}}$$

- where the subscripts correspond to the pattern being presented and a_m decreases with m .
- 15 After the full training set has been presented, the inverse Hessian matrix is given by $\mathbf{H}^{-1} = H_n^{-1}$. In algorithmic form, the Optimal Brain Surgeon method is:

begin initialize n_H, \mathbf{w}, θ

train a reasonably large network to minimum error

do compute \mathbf{H}^{-1} by Eqn. 1

- 20 $q^* \leftarrow \arg \min_q w_q^2 / (2[\mathbf{H}^{-1}]_{qq})$ (saliency L_q)

$$\mathbf{w} \leftarrow \mathbf{w} - \frac{w_q}{[H^{-1}]_{qq}} H^{-1} e_q \quad (\text{saliency } L_q)$$

until $J(\mathbf{w}) > \theta$

return \mathbf{w}

end

5 The Optimal Brain Damage method is computationally simpler because the calculation of the inverse Hessian matrix in line 3 is particularly simple for a diagonal matrix. The above algorithm terminates when the error is greater than a criterion initialized to be θ . Another approach is to change line 6 to terminate when the change in $J(\mathbf{w})$ due to elimination of a weight is greater than some criterion value.

10 In some embodiments, a back-propagation neural network (see, for example Abdi, 1994, "A neural network primer", J. Biol System. 2, 247-283) containing a single hidden layer of ten neurons (ten hidden units) found in EasyNN-Plus version 4.0g software package (Neural Planner Software Inc.) is used. In a specific example, parameter values within the EasyNN-Plus program are set as follows: a learning rate of 0.05, and a momentum of 0.2. In
15 some embodiments in which the EasyNN-Plus version 4.0g software package is used, "outlier" samples are identified by performing twenty independently-seeded trials involving 20,000 learning cycles each.

5.2.4.3. SUPPORT VECTOR MACHINE

20 In some embodiments of the present invention, support vector machines (SVMs) are used to classify subjects using expression profiles of marker genes described in the present invention. General description of SVM can be found in, for example, Cristianini and Shawe-Taylor, 2000, *An Introduction to Support Vector Machines*, Cambridge University Press, Cambridge; Boser *et al.*, 1992, "A training algorithm for optimal margin classifiers, in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, ACM
25 Press, Pittsburgh, PA, pp. 142-152; Vapnik, 1998, *Statistical Learning Theory*, Wiley, New York; Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc.; Hastie, 2001, *The Elements of Statistical Learning*, Springer, New York; and Furey *et al.*, 2000, *Bioinformatics* 16, 906-914. Applications of SVM in biological applications are described in Jaakkola *et al.*, *Proceedings of the 7th International Conference on Intelligent Systems for*

Molecular Biology, AAAI Press, Menlo Park, CA (1999); Brown *et al.*, *Proc. Natl. Acad. Sci.* 97(1):262-67 (2000); Zien *et al.*, *Bioinformatics*, 16(9):799-807 (2000); Furey *et al.*, *Bioinformatics*, 16(10):906-914 (2000)

In one approach, when a SVM is used, the gene expression data is standardized to have mean zero and unit variance and the members of a training population are randomly divided into a training set and a test set. For example, in one embodiment, two thirds of the members of the training population are placed in the training set and one third of the members of the training population are placed in the test set. The expression values for a selected set of genes of the present invention is used to train the SVM. Then the ability for the trained SVM to correctly classify members in the test set is determined. In some embodiments, this computation is performed several times for a given selected set of molecular markers. In each iteration of the computation, the members of the training population are randomly assigned to the training set and the test set. Then, the quality of the combination of molecular markers is taken as the average of each such iteration of the SVM computation.

Support vector machines map a given set of binary labeled training data to a high-dimensional feature space and separate the two classes of data with a maximum margin hyperplane. In general, this hyperplane corresponds to a nonlinear decision boundary in the input space. Let $X \in R_0 \subseteq \mathcal{R}^n$ be the input vectors, $y \in \{-1, +1\}$ be the labels, and $\phi: R_0 \rightarrow F$ be the mapping from input space to feature space. Then the SVM learning algorithm finds a hyperplane (w, b) such that the quantity

$$\gamma = \min_i y_i \{ \langle w, \phi(X_i) \rangle - b \}$$

is maximized, where the vector w has the same dimensionality as F , b is a real number, and γ is called the *margin*. The corresponding decision function is then

$$f(X) = \text{sign}(\langle w, \phi(X) \rangle - b)$$

This minimum occurs when

$$w = \sum_i \alpha_i y_i \phi(X_i)$$

where $\{\alpha_i\}$ are positive real numbers that maximize

$$\sum_i \alpha_i - \sum_{ij} \alpha_i \alpha_j y_i y_j \langle \phi(X_i), \phi(X_j) \rangle$$

subject to

$$\sum_i \alpha_i y_i = 0, \alpha_i > 0$$

5 The decision function can equivalently be expressed as

$$f(\mathbf{X}) = \text{sign}(\sum_i \alpha_i y_i \langle \phi(X_i), \phi(\mathbf{X}) \rangle - b)$$

From this equation it can be seen that the α_i associated with the training point \mathbf{X}_i expresses the strength with which that point is embedded in the final decision function. A remarkable property of this alternative representation is that only a subset of the points will be associated with a non-zero α_i . These points are called *support vectors* and are the points that lie closest to the separating hyperplane. The sparseness of the α vector has several computational and learning theoretic consequences. It is important to note that neither the learning algorithm nor the decision function needs to represent explicitly the image of points in the feature space, $\phi(\mathbf{X}_i)$, since both use only the dot products between such images, $\langle \phi(X_i), \phi(X_j) \rangle$. Hence, if one were given a function $K(\mathbf{X}, \mathbf{Y}) = \langle \phi(\mathbf{X}), \phi(\mathbf{Y}) \rangle$, one could learn and use the maximum margin hyperplane in the feature space without ever explicitly performing the mapping. For each continuous positive definite function $K(\mathbf{X}, \mathbf{Y})$ there exists a mapping ϕ such that $K(\mathbf{X}, \mathbf{Y}) = \langle \phi(\mathbf{X}), \phi(\mathbf{Y}) \rangle$ for all $\mathbf{X}, \mathbf{Y} \in R_0$ (Mercer's Theorem). The function $K(\mathbf{X}, \mathbf{Y})$ is called the *kernel function*. The use of a kernel function allows the support vector machine to operate efficiently in a nonlinear high-dimensional feature spaces without being adversely affected by the dimensionality of that space. Indeed, it is possible to work with feature spaces of infinite dimension. Moreover, Mercer's theorem makes it possible to learn in the feature space without even knowing ϕ and F . The matrix $K_{ij} = \langle \phi(\mathbf{X}_i), \phi(\mathbf{X}_j) \rangle$ is called the *kernel matrix*. Finally, note that the learning algorithm is a quadratic optimization problem that has only a global optimum. The absence of local minima is a significant difference from standard pattern recognition techniques such as neural

networks. For moderate sample sizes, the optimization problem can be solved with simple gradient descent techniques. In the presence of noise, the standard maximum margin algorithm described above can be subject to overfitting, and more sophisticated techniques should be used. This problem arises because the maximum margin algorithm always finds a perfectly consistent hypothesis and does not tolerate training error. Sometimes, however, it is necessary to trade some training accuracy for better predictive power. The need for tolerating training error has led to the development the soft-margin and the margin-distribution classifiers. One of these techniques replaces the kernel matrix in the training phase as follows:

$$K \leftarrow K + \lambda I$$

while still using the standard kernel function in the decision phase. By tuning λ , one can control the training error, and it is possible to prove that the risk of misclassifying unseen points can be decreased with a suitable choice of λ .

If instead of controlling the overall training error one wants to control the trade-off between false positives and false negatives, it is possible to modify K as follows:

$$K \leftarrow K + \lambda D$$

where D is a diagonal matrix whose entries are either d^+ or d^- , in locations corresponding to positive and negative examples. It is possible to prove that this technique is equivalent to controlling the size of the α_i in a way that depends on the size of the class, introducing a bias for larger α_i in the class with smaller d . This in turn corresponds to an asymmetric margin; i.e., the class with smaller d will be kept further away from the decision boundary. In some cases, the extreme imbalance of the two classes, along with the presence of noise, creates a situation in which points from the minority class can be easily mistaken for mislabelled points. Enforcing a strong bias against training errors in the minority class provides protection against such errors and forces the SVM to make the positive examples support vectors. Thus, choosing $d^+ = \frac{1}{n^+}$ and $d^- = \frac{1}{n^-}$ provides a heuristic way to automatically adjust the relative importance of the two classes, based on their respective cardinalities. This technique effectively controls the trade-off between sensitivity and specificity.

In the present invention, a linear kernel can be used. The similarity between two marker profiles X and Y can be the dot product $X \cdot Y$. In one embodiment, the kernel is

$$K(X, Y) = X \cdot Y + 1$$

In another embodiment, a kernel of degree d is used

5
$$K(X, Y) = (X \cdot Y + 1)^d, \text{ where } d \text{ can be either } 2, 3, \dots$$

In still another embodiment, a Gaussian kernel is used

$$K(X, Y) = \exp\left(\frac{-\|X - Y\|^2}{2\sigma^2}\right)$$

where σ is the width of the Gaussian.

5.2.4.4. LOGISTIC REGRESSION

10 In some embodiments, the classifier is based on a regression model, preferably a logistic regression model. Such a regression model includes a coefficient for each of the molecular markers in a selected set of molecular markers of the invention. In such embodiments, the coefficients for the regression model are computed using, for example, a maximum likelihood approach. In particular embodiments, molecular marker data from two
15 different clinical groups, e.g., good or poor prognosis, is used and the dependent variable is the clinical status of the patient for which molecular marker characteristic data are from.

Some embodiments of the present invention provide generalizations of the logistic regression model that handle multcategory (polychotomous) responses. Such embodiments can be used to discriminate an organism into one or three or more clinical groups, e.g., good,
20 intermediate, and poor prognosis. Such regression models use multcategory logit models that simultaneously refer to all pairs of categories, and describe the odds of response in one category instead of another. Once the model specifies logits for a certain $(J-1)$ pairs of categories, the rest are redundant. See, for example, Agresti, *An Introduction to Categorical Data Analysis*, John Wiley & Sons, Inc., 1996, New York, Chapter 8, which is hereby
25 incorporated by reference.

5.2.4.5. DISCRIMINANT ANALYSIS

Linear discriminant analysis (LDA) attempts to classify a subject into one of two categories based on certain object properties. In other words, LDA tests whether object attributes measured in an experiment predict categorization of the objects. LDA typically
30 requires continuous independent variables and a dichotomous categorical dependent variable. In the present invention, the expression values for the selected set of molecular markers of the

invention across a subset of the training population serve as the requisite continuous independent variables. The clinical group classification of each of the members of the training population serves as the dichotomous categorical dependent variable.

LDA seeks the linear combination of variables that maximizes the ratio of between-group variance and within-group variance by using the grouping information. Implicitly, the linear weights used by LDA depend on how the expression of a molecular marker across the training set separates in the two groups (*e.g.*, a group that has good prognosis and a group that have poor prognosis) and how this gene expression correlates with the expression of other genes. In some embodiments, LDA is applied to the data matrix of the N members in the training sample by K genes in a combination of genes described in the present invention. Then, the linear discriminant of each member of the training population is plotted. Ideally, those members of the training population representing a first subgroup (*e.g.* those subjects that good prognosis) will cluster into one range of linear discriminant values (*e.g.*, negative) and those member of the training population representing a second subgroup (*e.g.* those subjects that have poor prognosis) will cluster into a second range of linear discriminant values (*e.g.*, positive). The LDA is considered more successful when the separation between the clusters of discriminant values is larger. For more information on linear discriminant analysis, see Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc; and Hastie, 2001, *The Elements of Statistical Learning*, Springer, New York; Venables & Ripley, 1997, *Modern Applied Statistics with s-plus*, Springer, New York.

Quadratic discriminant analysis (QDA) takes the same input parameters and returns the same results as LDA. QDA uses quadratic equations, rather than linear equations, to produce results. LDA and QDA are interchangeable, and which to use is a matter of preference and/or availability of software to support the analysis. Logistic regression takes the same input parameters and returns the same results as LDA and QDA.

5.2.4.6. DECISION TREES

In some embodiments of the present invention, decision trees are used to classify patients using expression data for a selected set of molecular markers of the invention. Decision tree algorithms belong to the class of supervised learning algorithms. The aim of a decision tree is to induce a classifier (a tree) from real-world example data. This tree can be used to classify unseen examples which have not been used to derive the decision tree.

A decision tree is derived from training data. An example contains values for the different attributes and what class the example belongs. In one embodiment, the training data is expression data for a combination of genes described in the present invention across the training population.

5 The following algorithm describes a decision tree derivation:

Tree(Examples,Class,Attributes)

 Create a root node

 If all Examples have the same Class value, give the root this label

 Else if Attributes is empty label the root according to the most common value

10 Else begin

 Calculate the information gain for each attribute

 Select the attribute A with highest information gain and make this the root
attribute

 For each possible value, v, of this attribute

15 Add a new branch below the root, corresponding to $A = v$

 Let Examples(v) be those examples with $A = v$

 If Examples(v) is empty, make the new branch a leaf node labeled with
the most common value among Examples

 Else let the new branch be the tree created by

20 Tree(Examples(v),Class,Attributes - {A})

 end

A more detailed description of the calculation of information gain is shown in the following. If the possible classes v_i of the examples have probabilities $P(v_i)$ then the information content I of the actual answer is given by:

$$I(P(v_1), \dots, P(v_n)) = \sum_{i=1}^n -P(v_i) \log_2 P(v_i)$$

- 5 The I -value shows how much information we need in order to be able to describe the outcome of a classification for the specific dataset used. Supposing that the dataset contains p positive (*e.g.* has poor prognosis) and n negative (*e.g.* has good prognosis) examples (*e.g.* individuals), the information contained in a correct answer is:

$$I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) = -\frac{p}{p+n} \log_2 \frac{p}{p+n} - \frac{n}{p+n} \log_2 \frac{n}{p+n}$$

- 10 where \log_2 is the logarithm using base two. By testing single attributes the amount of information needed to make a correct classification can be reduced. The remainder for a specific attribute A (*e.g.* a gene) shows how much the information that is needed can be reduced.

$$\text{Remainder}(A) = \sum_{i=1}^v \frac{p_i + n_i}{p+n} I\left(\frac{p_i}{p_i + n_i}, \frac{n_i}{p_i + n_i}\right)$$

- 15 "v" is the number of unique attribute values for attribute A in a certain dataset, "i" is a certain attribute value, " p_i " is the number of examples for attribute A where the classification is positive (*e.g.* cancer), " n_i " is the number of examples for attribute A where the classification is negative (*e.g.* healthy).

- 20 The information gain of a specific attribute A is calculated as the difference between the information content for the classes and the remainder of attribute A :

$$\text{Gain}(A) = I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) - \text{Remainder}(A)$$

The information gain is used to evaluate how important the different attributes are for the classification (how well they split up the examples), and the attribute with the highest information.

5 In general there are a number of different decision tree algorithms, many of which are described in Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc. Decision tree algorithms often require consideration of feature processing, impurity measure, stopping criterion, and pruning. Specific decision tree algorithms include, but are not limited to classification and regression trees (CART), multivariate decision trees, ID3, and C4.5.

10 In one approach, when an exemplary embodiment of a decision tree is used, the gene expression data for a selected set of molecular markers of the invention across a training population is standardized to have mean zero and unit variance. The members of the training population are randomly divided into a training set and a test set. For example, in one embodiment, two thirds of the members of the training population are placed in the training set and one third of the members of the training population are placed in the test set. The
15 expression values for a select combination of genes described in the present invention is used to construct the decision tree. Then, the ability for the decision tree to correctly classify members in the test set is determined. In some embodiments, this computation is performed several times for a given combination of molecular markers. In each iteration of the computation, the members of the training population are randomly assigned to the training set
20 and the test set. Then, the quality of the combination of molecular markers is taken as the average of each such iteration of the decision tree computation.

5.2.4.7. CLUSTERING

In some embodiments, the expression values for a selected set of molecular markers of the invention are used to cluster a training set. For example, consider the case in which ten
25 genes described in the present invention are used. Each member m of the training population will have expression values for each of the ten genes. Such values from a member m in the training population define the vector:

X_{1m}	X_{2m}	X_{3m}	X_{4m}	X_{5m}	X_{6m}	X_{7m}	X_{8m}	X_{9m}	X_{10m}
----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------

where X_{im} is the expression level of the i^{th} gene in organism m . If there are m organisms in the training set, selection of i genes will define m vectors. Note that the methods of the present invention do not require that each the expression value of every single gene used in the vectors be represented in every single vector m . In other words, data from a subject in which one of the i^{th} genes is not found can still be used for clustering. In such instances, the missing expression value is assigned either a "zero" or some other normalized value. In some embodiments, prior to clustering, the gene expression values are normalized to have a mean value of zero and unit variance.

Those members of the training population that exhibit similar expression patterns across the training group will tend to cluster together. A particular combination of genes of the present invention is considered to be a good classifier in this aspect of the invention when the vectors cluster into the trait groups found in the training population. For instance, if the training population includes patients with good or poor prognosis, a clustering classifier will cluster the population into two groups, with each group uniquely representing either a good prognosis or a poor prognosis.

Clustering is described on pages 211-256 of Duda and Hart, *Pattern Classification and Scene Analysis*, 1973, John Wiley & Sons, Inc., New York. As described in Section 6.7 of Duda, the clustering problem is described as one of finding natural groupings in a dataset.

To identify natural groupings, two issues are addressed. First, a way to measure similarity (or dissimilarity) between two samples is determined. This metric (similarity measure) is used to ensure that the samples in one cluster are more like one another than they are to samples in other clusters. Second, a mechanism for partitioning the data into clusters using the similarity measure is determined.

Similarity measures are discussed in Section 6.7 of Duda, where it is stated that one way to begin a clustering investigation is to define a distance function and to compute the matrix of distances between all pairs of samples in a dataset. If distance is a good measure of similarity, then the distance between samples in the same cluster will be significantly less than the distance between samples in different clusters. However, as stated on page 215 of Duda, clustering does not require the use of a distance metric. For example, a nonmetric similarity function $s(x, x')$ can be used to compare two vectors x and x' . Conventionally, $s(x,$

x') is a symmetric function whose value is large when x and x' are somehow "similar". An example of a nonmetric similarity function $s(x, x')$ is provided on page 216 of Duda.

Once a method for measuring "similarity" or "dissimilarity" between points in a dataset has been selected, clustering requires a criterion function that measures the clustering quality of any partition of the data. Partitions of the data set that extremize the criterion function are used to cluster the data. See page 217 of Duda. Criterion functions are discussed in Section 6.8 of Duda.

More recently, Duda *et al.*, Pattern Classification, 2nd edition, John Wiley & Sons, Inc. New York, has been published. Pages 537-563 describe clustering in detail. More information on clustering techniques can be found in Kaufman and Rousseeuw, 1990, *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, New York, NY; Everitt, 1993, *Cluster analysis* (3d ed.), Wiley, New York, NY; and Backer, 1995, *Computer-Assisted Reasoning in Cluster Analysis*, Prentice Hall, Upper Saddle River, New Jersey. Particular exemplary clustering techniques that can be used in the present invention include, but are not limited to, hierarchical clustering (agglomerative clustering using nearest-neighbor algorithm, farthest-neighbor algorithm, the average linkage algorithm, the centroid algorithm, or the sum-of-squares algorithm), k-means clustering, fuzzy k-means clustering algorithm, and Jarvis-Patrick clustering.

5.2.4.8. PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) has been proposed to analyze gene expression data. Principal component analysis is a classical technique to reduce the dimensionality of a data set by transforming the data to a new set of variable (principal components) that summarize the features of the data. See, for example, Jolliffe, 1986, *Principal Component Analysis*, Springer, New York. Principal components (PCs) are uncorrelated and are ordered such that the k^{th} PC has the k^{th} largest variance among PCs. The k^{th} PC can be interpreted as the direction that maximizes the variation of the projections of the data points such that it is orthogonal to the first $k - 1$ PCs. The first few PCs capture most of the variation in the data set. In contrast, the last few PCs are often assumed to capture only the residual 'noise' in the data.

PCA can also be used to create a classifier in accordance with the present invention. In such an approach, vectors for a selected set of molecular markers of the invention can be

constructed in the same manner described for clustering above. In fact, the set of vectors, where each vector represents the expression values for the select genes from a particular member of the training population, can be considered a matrix. In some embodiments, this matrix is represented in a Free-Wilson method of qualitative binary description of monomers
5 (Kubinyi, 1990, *3D QSAR in drug design theory methods and applications*, Pergamon Press, Oxford, pp 589-638), and distributed in a maximally compressed space using PCA so that the first principal component (PC) captures the largest amount of variance information possible, the second principal component (PC) captures the second largest amount of all variance information, and so forth until all variance information in the matrix has been accounted for.

10 Then, each of the vectors (where each vector represents a member of the training population) is plotted. Many different types of plots are possible. In some embodiments, a one-dimensional plot is made. In this one-dimensional plot, the value for the first principal component from each of the members of the training population is plotted. In this form of plot, the expectation is that members of a first group (*e.g.* good prognosis patients) will
15 cluster in one range of first principal component values and members of a second group (*e.g.*, poor prognosis patients) will cluster in a second range of first principal component values.

In one example, the training population comprises two groups: a good prognosis group and a poor prognosis group. The first principal component is computed using the molecular marker expression values for the select genes of the present invention across the
20 entire training population data set. Then, each member of the training set is plotted as a function of the value for the first principal component. In this example, those members of the training population in which the first principal component is positive are the good prognosis patients and those members of the training population in which the first principal component is negative are poor prognosis patients.

25 In some embodiments, the members of the training population are plotted against more than one principal component. For example, in some embodiments, the members of the training population are plotted on a two-dimensional plot in which the first dimension is the first principal component and the second dimension is the second principal component. In such a two-dimensional plot, the expectation is that members of each subgroup represented in
30 the training population will cluster into discrete groups. For example, a first cluster of members in the two-dimensional plot will represent subjects with good prognosis, a second

cluster of members in the two-dimensional plot will represent subjects with poor prognosis, and so forth.

In some embodiments, the members of the training population are plotted against more than two principal components and a determination is made as to whether the members
5 of the training population are clustering into groups that each uniquely represents a subgroup found in the training population. In some embodiments, principal component analysis is performed by using the *R mva* package (Anderson, 1973, *Cluster Analysis for applications*, Academic Press, New York 1973; Gordon, *Classification*, Second Edition, Chapman and Hall, CRC, 1999.). Principal component analysis is further described in Duda, *Pattern*
10 *Classification*, Second Edition, 2001, John Wiley & Sons, Inc.

5.2.4.9. NEAREST NEIGHBOR CLASSIFIER ANALYSIS

Nearest neighbor classifiers are memory-based and require no model to be fit. Given a query point x_0 , the k training points $x_{(r)}$, $r = 1, \dots, k$ closest in distance to x_0 are identified and then the point x_0 is classified using the k nearest neighbors. Ties can be broken at random. In
15 some embodiments, Euclidean distance in feature space is used to determine distance as:

$$d_{(i)} = \|x_{(i)} - x_0\|.$$

Typically, when the nearest neighbor algorithm is used, the expression data used to compute the linear discriminant is standardized to have mean zero and variance 1. In the present invention, the members of the training population are randomly divided into a training set and
20 a test set. For example, in one embodiment, two thirds of the members of the training population are placed in the training set and one third of the members of the training population are placed in the test set. Profiles of a selected set of molecular markers of the invention represents the feature space into which members of the test set are plotted. Next, the ability of the training set to correctly characterize the members of the test set is computed.
25 In some embodiments, nearest neighbor computation is performed several times for a given combination of genes of the present invention. In each iteration of the computation, the members of the training population are randomly assigned to the training set and the test set. Then, the quality of the combination of genes is taken as the average of each such iteration of the nearest neighbor computation.

The nearest neighbor rule can be refined to deal with issues of unequal class priors, differential misclassification costs, and feature selection. Many of these refinements involve some form of weighted voting for the neighbors. For more information on nearest neighbor analysis, see Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc; and Hastie, 2001, *The Elements of Statistical Learning*, Springer, New York.

5.2.4.10. EVOLUTIONARY METHODS

Inspired by the process of biological evolution, evolutionary methods of classifier design employ a stochastic search for an optimal classifier. In broad overview, such methods create several classifiers - a population - from measurements of gene products of the present invention. Each classifier varies somewhat from the other. Next, the classifiers are scored on expression data across the training population. In keeping with the analogy with biological evolution, the resulting (scalar) score is sometimes called the fitness. The classifiers are ranked according to their score and the best classifiers are retained (some portion of the total population of classifiers). Again, in keeping with biological terminology, this is called survival of the fittest. The classifiers are stochastically altered in the next generation - the children or offspring. Some offspring classifiers will have higher scores than their parent in the previous generation, some will have lower scores. The overall process is then repeated for the subsequent generation: The classifiers are scored and the best ones are retained, randomly altered to give yet another generation, and so on. In part, because of the ranking, each generation has, on average, a slightly higher score than the previous one. The process is halted when the single best classifier in a generation has a score that exceeds a desired criterion value. More information on evolutionary methods is found in, for example, Duda, *Pattern Classification*, Second Edition, 2001, John Wiley & Sons, Inc.

5.2.4.11. BAGGING, BOOSTING AND THE RANDOM SUBSPACE METHOD

Bagging, boosting and the random subspace method are combining techniques that can be used to improve weak classifiers. These techniques are designed for, and usually applied to, decision trees. In addition, Skurichina and Duin provide evidence to suggest that such techniques can also be useful in linear discriminant analysis.

In bagging, one samples the training set, generating random independent bootstrap replicates, constructs the classifier on each of these, and aggregates them by a simple majority vote in the final decision rule. See, for example, Breiman, 1996, *Machine Learning*

24, 123-140; and Efron & Tibshirani, *An Introduction to Bootstrap*, Chapman & Hall, New York, 1993.

In boosting, classifiers are constructed on weighted versions of the training set, which are dependent on previous classification results. Initially, all objects have equal weights, and the first classifier is constructed on this data set. Then, weights are changed according to the performance of the classifier. Erroneously classified objects (molecular markers in the data set) get larger weights, and the next classifier is boosted on the reweighted training set. In this way, a sequence of training sets and classifiers is obtained, which is then combined by simple majority voting or by weighted majority voting in the final decision. See, for example, Freund & Schapire, "Experiments with a new boosting algorithm," Proceedings 13th International Conference on Machine Learning, 1996, 148-156.

To illustrate boosting, consider the case where there are two phenotypic groups exhibited by the population under study, phenotype 1 (e.g., good prognosis patients), and phenotype 2 (e.g., poor prognosis patients). Given a vector of molecular markers X , a classifier $G(X)$ produces a prediction taking one of the type values in the two value set: {phenotype 1, phenotype 2}. The error rate on the training sample is

$$\overline{\text{err}} = \frac{1}{N} \sum_{i=1}^N I(y_i \neq G(x_i))$$

where N is the number of subjects in the training set (the sum total of the subjects that have either phenotype 1 or phenotype 2).

A weak classifier is one whose error rate is only slightly better than random guessing. In the boosting algorithm, the weak classification algorithm is repeatedly applied to modified versions of the data, thereby producing a sequence of weak classifiers $G_m(x)$, $m = 1, 2, \dots, M$. The predictions from all of the classifiers in this sequence are then combined through a weighted majority vote to produce the final prediction:

$$G(x) = \text{sign} \left(\sum_{m=1}^M \alpha_m G_m(x) \right)$$

Here $\alpha_1, \alpha_2, \dots, \alpha_M$ are computed by the boosting algorithm and their purpose is to weigh the contribution of each respective $G_m(x)$. Their effect is to give higher influence to the more accurate classifiers in the sequence.

5 The data modifications at each boosting step consist of applying weights w_1, w_2, \dots, w_n to each of the training observations (x_i, y_i) , $i = 1, 2, \dots, N$. Initially all the weights are set to $w_i = 1/N$, so that the first step simply trains the classifier on the data in the usual manner. For each successive iteration $m = 2, 3, \dots, M$ the observation weights are individually modified and the classification algorithm is reapplied to the weighted observations. At stem
10 m , those observations that were misclassified by the classifier $G_{m-1}(x)$ induced at the previous step have their weights increased, whereas the weights are decreased for those that were classified correctly. Thus as iterations proceed, observations that are difficult to correctly classify receive ever-increasing influence. Each successive classifier is thereby forced to concentrate on those training observations that are missed by previous ones in the sequence.

15 The exemplary boosting algorithm is summarized as follows:

1. Initialize the observation weights $w_i = 1/N$, $i = 1, 2, \dots, N$.

2. For $m = 1$ to M :

20 (a) Fit a classifier $G_m(x)$ to the training set using weights w_i .

(b) Compute

$$\text{err}_m = \frac{\sum_{i=1}^N w_i I(y_i \neq G_m(x_i))}{\sum_{i=1}^N w_i}$$

(c) Compute $\alpha_m = \log((1 - \text{err}_m) / \text{err}_m)$.

(d) Set $w_i \leftarrow w_i \cdot \exp[\alpha_m \cdot I(y_i \neq G_m(x_i))]$, $i = 1, 2, \dots, N$.

3. Output $G(x) = \text{sign} \left[\sum_{m=1}^M \alpha_m G_m(x) \right]$

In the algorithm, the current classifier $G_m(x)$ is induced on the weighted observations at line 5
 2a. The resulting weighted error rate is computed at line 2b. Line 2c calculates the weight α_m
 given to $G_m(x)$ in producing the final classifier $G(x)$ (line 3). The individual weights of each
 of the observations are updated for the next iteration at line 2d. Observations misclassified
 by $G_m(x)$ have their weights scaled by a factor $\exp(\alpha_m)$, increasing their relative influence for
 inducing the next classifier $G_{m+1}(x)$ in the sequence. In some embodiments, modifications of
 10 the Freund and Schapire, 1997, Journal of Computer and System Sciences 55, pp. 119-139,
 boosting method are used. See, for example, Hasti *et al.*, *The Elements of Statistical
 Learning*, 2001, Springer, New York, Chapter 10. In some embodiments, boosting or
 adaptive boosting methods are used.

In some embodiments, modifications of Freund and Schapire, 1997, Journal of
 15 Computer and System Sciences 55, pp. 119-139, are used. For example, in some
 embodiments, feature preselection is performed using a technique such as the nonparametric
 scoring methods of Park *et al.*, 2002, Pac. Symp. Biocomput. 6, 52-63. Feature preselection
 is a form of dimensionality reduction in which the genes that discriminate between
 classifications the best are selected for use in the classifier. Then, the LogitBoost procedure
 20 introduced by Friedman *et al.*, 2000, Ann Stat 28, 337-407 is used rather than the boosting
 procedure of Freund and Schapire. In some embodiments, the boosting and other
 classification methods of Ben-Dor *et al.*, 2000, Journal of Computational Biology 7, 559-583
 are used in the present invention. In some embodiments, the boosting and other classification
 methods of Freund and Schapire, 1997, Journal of Computer and System Sciences 55, 119-
 25 139, are used.

In the random subspace method, classifiers are constructed in random subspaces of
 the data feature space. These classifiers are usually combined by simple majority voting in
 the final decision rule. See, for example, Ho, "The Random subspace method for
 constructing decision forests," IEEE Trans Pattern Analysis and Machine Intelligence, 1998;
 30 20(8): 832-844.

5.2.4.12. OTHER ALGORITHMS

The pattern classification and statistical techniques described above are merely examples of the types of models that can be used to construct a model for classification. Moreover, combinations of the techniques described above can be used. Some combinations, such as the use of the combination of decision trees and boosting, have been described. However, many other combinations are possible. In addition, in other techniques in the art such as Projection Pursuit and Weighted Voting can be used to construct a classifier.

5.3. DETERMINATION OF MARKER GENE EXPRESSION LEVELS

5.3.1. METHODS

10 The expression levels of the marker genes in a sample may be determined by any means known in the art. The expression level may be determined by isolating and determining the level (*i.e.*, amount) of nucleic acid transcribed from each marker gene. Alternatively, or additionally, the level of specific proteins encoded by a marker gene may be determined.

15 The level of expression of specific marker genes can be accomplished by determining the amount of mRNA, or polynucleotides derived therefrom, present in a sample. Any method for determining RNA levels can be used. For example, RNA is isolated from a sample and separated on an agarose gel. The separated RNA is then transferred to a solid support, such as a filter. Nucleic acid probes representing one or more markers are then
20 hybridized to the filter by northern hybridization, and the amount of marker-derived RNA is determined. Such determination can be visual, or machine-aided, for example, by use of a densitometer. Another method of determining RNA levels is by use of a dot-blot or a slot-blot. In this method, RNA, or nucleic acid derived therefrom, from a sample is labeled. The RNA or nucleic acid derived therefrom is then hybridized to a filter containing
25 oligonucleotides derived from one or more marker genes, wherein the oligonucleotides are placed upon the filter at discrete, easily-identifiable locations. Hybridization, or lack thereof, of the labeled RNA to the filter-bound oligonucleotides is determined visually or by densitometer. Polynucleotides can be labeled using a radiolabel or a fluorescent (*i.e.*, visible) label.

30 These examples are not intended to be limiting; other methods of determining RNA abundance are known in the art.

The level of expression of particular marker genes may also be assessed by determining the level of the specific protein expressed from the marker genes. This can be accomplished, for example, by separation of proteins from a sample on a polyacrylamide gel, followed by identification of specific marker-derived proteins using antibodies in a western blot. Alternatively, proteins can be separated by two-dimensional gel electrophoresis systems. Two-dimensional gel electrophoresis is well-known in the art and typically involves isoelectric focusing along a first dimension followed by SDS-PAGE electrophoresis along a second dimension. *See, e.g., Hames et al., 1990, GEL ELECTROPHORESIS OF PROTEINS: A PRACTICAL APPROACH, IRL Press, New York; Shevchenko et al., Proc. Nat'l Acad. Sci. USA* 93:1440-1445 (1996); Sagliocco *et al.*, *Yeast* 12:1519-1533 (1996); Lander, *Science* 274:536-539 (1996). The resulting electropherograms can be analyzed by numerous techniques, including mass spectrometric techniques, western blotting and immunoblot analysis using polyclonal and monoclonal antibodies.

Alternatively, marker-derived protein levels can be determined by constructing an antibody microarray in which binding sites comprise immobilized, preferably monoclonal, antibodies specific to a plurality of protein species encoded by the cell genome. Preferably, antibodies are present for a substantial fraction of the marker-derived proteins of interest. Methods for making monoclonal antibodies are well known (*see, e.g., Harlow and Lane, 1988, ANTIBODIES: A LABORATORY MANUAL, Cold Spring Harbor, New York, which is incorporated in its entirety for all purposes*). In one embodiment, monoclonal antibodies are raised against synthetic peptide fragments designed based on genomic sequence of the cell. With such an antibody array, proteins from the cell are contacted to the array, and their binding is assayed with assays known in the art. Generally, the expression, and the level of expression, of proteins of diagnostic or prognostic interest can be detected through immunohistochemical staining of tissue slices or sections.

Finally, expression of marker genes in a number of tissue specimens may be characterized using a "tissue array" (Kononen *et al.*, *Nat. Med* 4(7):844-7 (1998)). In a tissue array, multiple tissue samples are assessed on the same microarray. The arrays allow *in situ* detection of RNA and protein levels; consecutive sections allow the analysis of multiple samples simultaneously.

5.3.2. MICROARRAYS

In preferred embodiments, polynucleotide microarrays are used to measure expression so that the expression status of each of the markers above is assessed simultaneously.

Generally, microarrays according to the invention comprise a plurality of markers

informative for prognosis, or outcome determination, for a particular disease or condition,

5 and, in particular, for individuals having specific combinations of genotypic or phenotypic characteristics of the disease or condition (*i.e.*, that are prognosis-informative for a particular patient subset).

The microarrays of the invention preferably comprise at least 2, 3, 4, 5, 7, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200 or more of markers, or all of the markers, or any

10 combination of markers, identified as prognosis-informative within a patient subset. The actual number of informative markers the microarray comprises will vary depending upon the particular condition of interest, the number of markers identified, and, optionally, the number of informative markers found to result in the least Type I error, Type II error, or Type I and Type II error in determination of prognosis. As used herein, "Type I error" means a false
15 positive and "Type II error" means a false negative; in the example of prognosis of breast cancer, Type I error is the mis-characterization of an individual with a good prognosis as having a poor prognosis, and Type II error is the mis-characterization of an individual with a poor prognosis as having a good prognosis.

In specific embodiments, the invention provides polynucleotide arrays in which the
20 prognosis markers identified for a particular patient subset comprise at least 50%, 60%, 70%, 80%, 85%, 90%, 95% or 98% of the probes on said array. In another specific embodiment, the microarray comprises a plurality of probes, wherein said plurality of probes comprise probes complementary and hybridizable to at least 75% of the prognosis-informative markers identified for a particular patient subset. Microarrays of the invention, of course, may

25 comprise probes complementary and hybridizable to prognosis-informative markers for a plurality of the patient subsets, or for each patient subset, identified for a particular condition. In another embodiment, therefore, the microarray of the invention comprises a plurality of probes complementary and hybridizable to at least 75% of the prognosis-informative markers identified for each patient subset identified for the condition of interest, and wherein said
30 probes, in total, are at least 50% of the probes on said microarray.

In yet another specific embodiment, microarrays that are used in the methods disclosed herein optionally comprise markers additional to at least some of the markers

identified by the methods disclosed elsewhere herein. For example, in a specific embodiment, the microarray is a screening or scanning array as described in Altschuler *et al.*, International Publication WO 02/18646, published March 7, 2002 and Scherer *et al.*, International Publication WO 02/16650, published February 28, 2002. The scanning and screening arrays comprise regularly-spaced, positionally-addressable probes derived from genomic nucleic acid sequence, both expressed and unexpressed. Such arrays may comprise probes corresponding to a subset of, or all of, the markers identified for the patient subset(s) for the condition of interest, and can be used to monitor marker expression in the same way as a microarray containing only prognosis-informative markers otherwise identified.

In yet another specific embodiment, the microarray is a commercially-available cDNA microarray that comprises at least five markers identified by the methods described herein. Preferably, a commercially-available cDNA microarray comprises all of the markers identified by the methods described herein as being informative for a patient subset for a particular condition. However, such a microarray may comprise at least 5, 10, 15 or 25 of such markers, up to the maximum number of markers identified.

In an embodiment specific to breast cancer, the invention provides for oligonucleotide or cDNA arrays comprising probes hybridizable to the genes corresponding to each of the marker sets described above (*i.e.*, markers informative for ER⁻, sporadic individuals, markers informative for ER⁻, *BRCA1* individuals, markers informative for ER⁺, ER/AGE high individuals, markers informative for ER⁺, ER/AGE low, LN⁺ individuals, and markers informative for ER⁺, ER/AGE low, LN⁻ individuals, as shown in Tables 1-5). Any of the microarrays described herein may be provided in a sealed container in a kit.

The invention provides microarrays containing probes useful for the prognosis of any breast cancer patient, or for breast cancer patients classified into one of a plurality of patient subsets. In particular, the invention provides polynucleotide arrays comprising probes to a subset or subsets of at least 5, 10, 15, 20, 25 or more of the genetic markers, or up to the full set of markers, in any of Tables 1-5, which distinguish between patients with good and poor prognosis. In certain embodiments, therefore, the invention provides microarrays comprising probes for a plurality of the genes for which markers are listed in Tables 1, 2, 3, 4 or 5. In a specific embodiment, the microarray of the invention comprises 1, 2, 3, 4, 5 or 10 of the markers in Table 1, at least five of the markers in Table 2; 1, 2, 3, 4, 5 or 10 of the markers in Table 3; 1, 2, 3, 4, 5 or 10 of the markers in Table 4; or 1, 2, 3, 4, 5 or 10 of the markers in

Table 1. In other embodiments, the microarray comprises probes for 1, 2, 3, 4, 5, or 10 of the markers shown in any two, three or four of Tables 1-5, or all of Tables 1-5. In other embodiments, the microarray of the invention contains each of the markers in Table 1, Table 2, Table 3, Table 4, or Table 5. In another embodiment, the microarray contains all of the markers shown in Tables 1-5. In specific embodiments, the array comprises probes derived only from the markers listed in Table 1, Table 2, Table 3, Table 4, or Table 5; probes derived from any two of Tables 1-5; any three of Tables 1-5; any four of Tables 1-5; or all of Tables 1-5.

In other embodiments, the array comprises a plurality of probes derived from markers listed in any of Tables 1-5 in combination with a plurality of other probes, derived from markers not listed in any of Tables 1-5, that are identified as informative for the prognosis of breast cancer.

In specific embodiments, the invention provides polynucleotide arrays in which the breast cancer prognosis markers described herein in Tables 1, 2, 3, 4 and/or 5 comprise at least 50%, 60%, 70%, 80%, 85%, 90%, 95% or 98% of the probes on said array. In another specific embodiment, the microarray comprises a plurality of probes, wherein said plurality of probes comprise probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 1; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 2; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 3; probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 4; and probes complementary and hybridizable to at least 75% of the genes for which markers are listed in Table 5, wherein said probes, in total, are at least 50% of the probes on said microarray.

In yet another specific embodiment, microarrays that are used in the methods disclosed herein optionally comprise markers additional to at least some of the markers listed in Tables 1-5. For example, in a specific embodiment, the microarray is a screening or scanning array as described in Altschuler *et al.*, International Publication WO 02/18646, published March 7, 2002 and Scherer *et al.*, International Publication WO 02/16650, published February 28, 2002. The scanning and screening arrays comprise regularly-spaced, positionally-addressable probes derived from genomic nucleic acid sequence, both expressed and unexpressed. Such arrays may comprise probes corresponding to a subset of, or all of,

the markers listed in Tables 1-5, or a subset thereof as described above, and can be used to monitor marker expression in the same way as a microarray containing only markers listed in Tables 1-5.

In yet another specific embodiment, the microarray is a commercially-available cDNA microarray that comprises at least five of the markers listed in Tables 1-5. Preferably, a commercially-available cDNA microarray comprises all of the markers listed in Tables 1-5. However, such a microarray may comprise at least 5, 10, 15 or 25 of the markers in any of Tables 1-5, up to the maximum number of markers in a Table, and may comprise all of the markers in any one of Tables 1-5, and a subset of another of Tables 1-5, or subsets of each as described above. In a specific embodiment of the microarrays used in the methods disclosed herein, the markers that are all or a portion of Tables 1-5 make up at least 50%, 60%, 70%, 80%, 90%, 95% or 98% of the probes on the microarray.

General methods pertaining to the construction of microarrays comprising the marker sets and/or subsets above are described in the following sections.

5.3.2.1. CONSTRUCTION OF MICROARRAYS

Microarrays are prepared by selecting probes which comprise a polynucleotide sequence, and then immobilizing such probes to a solid support or surface. For example, the probes may comprise DNA sequences, RNA sequences, or copolymer sequences of DNA and RNA. The polynucleotide sequences of the probes may also comprise DNA and/or RNA analogues, or combinations thereof. For example, the polynucleotide sequences of the probes may be full or partial fragments of genomic DNA. The polynucleotide sequences of the probes may also be synthesized nucleotide sequences, such as synthetic oligonucleotide sequences. The probe sequences can be synthesized either enzymatically *in vivo*, enzymatically *in vitro* (e.g., by PCR), or non-enzymatically *in vitro*.

The probe or probes used in the methods of the invention are preferably immobilized to a solid support which may be either porous or non-porous. For example, the probes of the invention may be polynucleotide sequences which are attached to a nitrocellulose or nylon membrane or filter covalently at either the 3' or the 5' end of the polynucleotide. Such hybridization probes are well known in the art (see, e.g., Sambrook *et al.*, MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989). Alternatively, the solid support or surface may be a

glass or plastic surface. In a particularly preferred embodiment, hybridization levels are measured to microarrays of probes consisting of a solid phase on the surface of which are immobilized a population of polynucleotides, such as a population of DNA or DNA mimics, or, alternatively, a population of RNA or RNA mimics. The solid phase may be a nonporous
5 or, optionally, a porous material such as a gel.

In preferred embodiments, a microarray comprises a support or surface with an ordered array of binding (*e.g.*, hybridization) sites or "probes" each representing one of the markers described herein. Preferably the microarrays are addressable arrays, and more preferably positionally addressable arrays. More specifically, each probe of the array is
10 preferably located at a known, predetermined position on the solid support such that the identity (*i.e.*, the sequence) of each probe can be determined from its position in the array (*i.e.*, on the support or surface). In preferred embodiments, each probe is covalently attached to the solid support at a single site.

Microarrays can be made in a number of ways, of which several are described below.
15 However produced, microarrays share certain characteristics. The arrays are reproducible, allowing multiple copies of a given array to be produced and easily compared with each other. Preferably, microarrays are made from materials that are stable under binding (*e.g.*, nucleic acid hybridization) conditions. The microarrays are preferably small, *e.g.*, between 1 cm² and 25 cm², between 12 cm² and 13 cm², or 3 cm². However, larger arrays are also
20 contemplated and may be preferable, *e.g.*, for use in screening arrays. Preferably, a given binding site or unique set of binding sites in the microarray will specifically bind (*e.g.*, hybridize) to the product of a single gene in a cell (*e.g.*, to a specific mRNA, or to a specific cDNA derived therefrom). However, in general, other related or similar sequences will cross hybridize to a given binding site.

25 The microarrays of the present invention include one or more test probes, each of which has a polynucleotide sequence that is complementary to a subsequence of RNA or DNA to be detected. Preferably, the position of each probe on the solid surface is known. Indeed, the microarrays are preferably positionally addressable arrays. Specifically, each probe of the array is preferably located at a known, predetermined position on the solid
30 support such that the identity (*i.e.*, the sequence) of each probe can be determined from its position on the array (*i.e.*, on the support or surface).

According to the invention, the microarray is an array (*i.e.*, a matrix) in which each position represents one of the markers described herein. For example, each position can contain a DNA or DNA analogue based on genomic DNA to which a particular RNA or cDNA transcribed from that genetic marker can specifically hybridize. The DNA or DNA analogue can be, *e.g.*, a synthetic oligomer or a gene fragment. In one embodiment, probes representing each of the markers is present on the array. In a preferred embodiment, the array comprises probes for each of the markers listed in Tables 1-5.

5.3.2.2. PREPARING PROBES FOR MICROARRAYS

As noted above, the “probe” to which a particular polynucleotide molecule specifically hybridizes according to the invention contains a complementary genomic polynucleotide sequence. The probes of the microarray preferably consist of nucleotide sequences of no more than 1,000 nucleotides. In some embodiments, the probes of the array consist of nucleotide sequences of 10 to 1,000 nucleotides. In a preferred embodiment, the nucleotide sequences of the probes are in the range of 10-200 nucleotides in length and are genomic sequences of a species of organism, such that a plurality of different probes is present, with sequences complementary and thus capable of hybridizing to the genome of such a species of organism, sequentially tiled across all or a portion of such genome. In other specific embodiments, the probes are in the range of 10-30 nucleotides in length, in the range of 10-40 nucleotides in length, in the range of 20-50 nucleotides in length, in the range of 40-80 nucleotides in length, in the range of 50-150 nucleotides in length, in the range of 80-120 nucleotides in length, and most preferably are 60 nucleotides in length.

The probes may comprise DNA or DNA “mimics” (*e.g.*, derivatives and analogues) corresponding to a portion of an organism’s genome. In another embodiment, the probes of the microarray are complementary RNA or RNA mimics. DNA mimics are polymers composed of subunits capable of specific, Watson-Crick-like hybridization with DNA, or of specific hybridization with RNA. The nucleic acids can be modified at the base moiety, at the sugar moiety, or at the phosphate backbone. Exemplary DNA mimics include, *e.g.*, phosphorothioates.

DNA can be obtained, *e.g.*, by polymerase chain reaction (PCR) amplification of genomic DNA or cloned sequences. PCR primers are preferably chosen based on a known sequence of the genome that will result in amplification of specific fragments of genomic

DNA. Computer programs that are well known in the art are useful in the design of primers with the required specificity and optimal amplification properties, such as *Oligo* version 5.0 (National Biosciences). Typically each probe on the microarray will be between 10 bases and 50,000 bases, usually between 300 bases and 1,000 bases in length. PCR methods are well known in the art, and are described, for example, in Innis *et al.*, eds., PCR PROTOCOLS: A GUIDE TO METHODS AND APPLICATIONS, Academic Press Inc., San Diego, CA (1990). It will be apparent to one skilled in the art that controlled robotic systems are useful for isolating and amplifying nucleic acids.

An alternative, preferred means for generating the polynucleotide probes of the microarray is by synthesis of synthetic polynucleotides or oligonucleotides, *e.g.*, using N-phosphonate or phosphoramidite chemistries (Froehler *et al.*, *Nucleic Acid Res.* 14:5399-5407 (1986); McBride *et al.*, *Tetrahedron Lett.* 24:246-248 (1983)). Synthetic sequences are typically between about 10 and about 500 bases in length, more typically between about 20 and about 100 bases, and most preferably between about 40 and about 70 bases in length. In some embodiments, synthetic nucleic acids include non-natural bases, such as, but by no means limited to, inosine. As noted above, nucleic acid analogues may be used as binding sites for hybridization. An example of a suitable nucleic acid analogue is peptide nucleic acid (*see, e.g.*, Egholm *et al.*, *Nature* 363:566-568 (1993); U.S. Patent No. 5,539,083).

Probes are preferably selected using an algorithm that takes into account binding energies, base composition, sequence complexity, cross-hybridization binding energies, and secondary structure. *See* Friend *et al.*, International Patent Publication WO 01/05935, published January 25, 2001; Hughes *et al.*, *Nat. Biotech.* 19:342-7 (2001).

A skilled artisan will also appreciate that positive control probes, *e.g.*, probes known to be complementary and hybridizable to sequences in the target polynucleotide molecules, and negative control probes, *e.g.*, probes known to not be complementary and hybridizable to sequences in the target polynucleotide molecules, should be included on the array. In one embodiment, positive controls are synthesized along the perimeter of the array. In another embodiment, positive controls are synthesized in diagonal stripes across the array. In still another embodiment, the reverse complement for each probe is synthesized next to the position of the probe to serve as a negative control. In yet another embodiment, sequences from other species of organism are used as negative controls or as "spike-in" controls.

5.3.2.3. ATTACHING PROBES TO THE SOLID SURFACE

The probes are attached to a solid support or surface, which may be made, *e.g.*, from glass, plastic (*e.g.*, polypropylene, nylon), polyacrylamide, nitrocellulose, gel, or other porous or nonporous material. A preferred method for attaching the nucleic acids to a surface is by printing on glass plates, as is described generally by Schena *et al.*, *Science* 270:467-470 (1995). This method is especially useful for preparing microarrays of cDNA (See also, DeRisi *et al.*, *Nature Genetics* 14:457-460 (1996); Shalon *et al.*, *Genome Res.* 6 :639-645 (1996); and Schena *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 93:10539-11286 (1995)).

A second preferred method for making microarrays is by making high-density oligonucleotide arrays. Techniques are known for producing arrays containing thousands of oligonucleotides complementary to defined sequences, at defined locations on a surface using photolithographic techniques for synthesis *in situ* (see, Fodor *et al.*, 1991, *Science* 251:767-773; Pease *et al.*, 1994, *Proc. Natl. Acad. Sci. U.S.A.* 91:5022-5026; Lockhart *et al.*, 1996, *Nature Biotechnology* 14:1675; U.S. Patent Nos. 5,578,832; 5,556,752; and 5,510,270) or other methods for rapid synthesis and deposition of defined oligonucleotides (Blanchard *et al.*, *Biosensors & Bioelectronics* 11:687-690). When these methods are used, oligonucleotides (*e.g.*, 60-mers) of known sequence are synthesized directly on a surface such as a derivatized glass slide. Usually, the array produced is redundant, with several oligonucleotide molecules per RNA.

Other methods for making microarrays, *e.g.*, by masking (Maskos and Southern, 1992, *Nuc. Acids. Res.* 20:1679-1684), may also be used. In principle, and as noted *supra*, any type of array, for example, dot blots on a nylon hybridization membrane (see Sambrook *et al.*, *MOLECULAR CLONING - A LABORATORY MANUAL* (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989)) could be used. However, as will be recognized by those skilled in the art, very small arrays will frequently be preferred because hybridization volumes will be smaller.

In one embodiment, the arrays of the present invention are prepared by synthesizing polynucleotide probes on a support. In such an embodiment, polynucleotide probes are attached to the support covalently at either the 3' or the 5' end of the polynucleotide.

In a particularly preferred embodiment, microarrays of the invention are manufactured by means of an ink jet printing device for oligonucleotide synthesis, *e.g.*, using

the methods and systems described by Blanchard in U.S. Pat. No. 6,028,189; Blanchard et al., 1996, *Biosensors and Bioelectronics* 11:687-690; Blanchard, 1998, in *Synthetic DNA Arrays in Genetic Engineering*, Vol. 20, J.K. Setlow, Ed., Plenum Press, New York at pages 111-123. Specifically, the oligonucleotide probes in such microarrays are preferably
5 synthesized in arrays, e.g., on a glass slide, by serially depositing individual nucleotide bases in "microdroplets" of a high surface tension solvent such as propylene carbonate. The microdroplets have small volumes (e.g., 100 pL or less, more preferably 50 pL or less) and are separated from each other on the microarray (e.g., by hydrophobic domains) to form circular surface tension wells which define the locations of the array elements (i.e., the
10 different probes). Microarrays manufactured by this ink-jet method are typically of high density, preferably having a density of at least about 2,500 different probes per 1 cm². The polynucleotide probes are attached to the support covalently at either the 3' or the 5' end of the polynucleotide.

5.3.2.4. TARGET POLYNUCLEOTIDE MOLECULES

15 The polynucleotide molecules which may be analyzed by the present invention (the "target polynucleotide molecules") may be from any clinically relevant source, but are expressed RNA or a nucleic acid derived therefrom (e.g., cDNA or amplified RNA derived from cDNA that incorporates an RNA polymerase promoter), including naturally occurring nucleic acid molecules, as well as synthetic nucleic acid molecules. In one embodiment, the
20 target polynucleotide molecules comprise RNA, including, but by no means limited to, total cellular RNA, poly(A)⁺ messenger RNA (mRNA) or fraction thereof, cytoplasmic mRNA, or RNA transcribed from cDNA (i.e., cRNA; see, e.g., Linsley & Schelter, U.S. Patent Application No. 09/411,074, filed October 4, 1999, or U.S. Patent Nos. 5,545,522, 5,891,636, or 5,716,785). Methods for preparing total and poly(A)⁺ RNA are well known in the art, and
25 are described generally, e.g., in Sambrook *et al.*, *MOLECULAR CLONING - A LABORATORY MANUAL* (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989). In one embodiment, RNA is extracted from cells of the various types of interest in this invention using guanidinium thiocyanate lysis followed by CsCl centrifugation (Chirgwin *et al.*, 1979, *Biochemistry* 18:5294-5299). In another embodiment, total RNA is
30 extracted using a silica gel-based column, commercially available examples of which include RNeasy (Qiagen, Valencia, California) and StrataPrep (Stratagene, La Jolla, California). In an alternative embodiment, which is preferred for *S. cerevisiae*, RNA is extracted from cells

using phenol and chloroform, as described in Ausubel *et al.*, eds., 1989, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Vol. III, Green Publishing Associates, Inc., John Wiley & Sons, Inc., New York, at pp. 13.12.1-13.12.5). Poly(A)⁺ RNA can be selected, *e.g.*, by selection with oligo-dT cellulose or, alternatively, by oligo-dT primed reverse transcription of
5 total cellular RNA. In one embodiment, RNA can be fragmented by methods known in the art, *e.g.*, by incubation with ZnCl₂, to generate fragments of RNA. In another embodiment, the polynucleotide molecules analyzed by the invention comprise cDNA, or PCR products of amplified RNA or cDNA.

In one embodiment, total RNA, mRNA, or nucleic acids derived therefrom, is isolated
10 from a sample taken from a person afflicted with breast cancer. Target polynucleotide molecules that are poorly expressed in particular cells may be enriched using normalization techniques (Bonaldi *et al.*, 1996, *Genome Res.* 6:791-806).

As described above, the target polynucleotides are detectably labeled at one or more nucleotides. Any method known in the art may be used to detectably label the target
15 polynucleotides. Preferably, this labeling incorporates the label uniformly along the length of the RNA, and more preferably, the labeling is carried out at a high degree of efficiency. One embodiment for this labeling uses oligo-dT primed reverse transcription to incorporate the label; however, conventional methods of this method are biased toward generating 3' end fragments. Thus, in a preferred embodiment, random primers (*e.g.*, 9-mers) are used in
20 reverse transcription to uniformly incorporate labeled nucleotides over the full length of the target polynucleotides. Alternatively, random primers may be used in conjunction with PCR methods or T7 promoter-based *in vitro* transcription methods in order to amplify the target polynucleotides.

In a preferred embodiment, the detectable label is a luminescent label. For example,
25 fluorescent labels, bioluminescent labels, chemiluminescent labels, and colorimetric labels may be used in the present invention. In a highly preferred embodiment, the label is a fluorescent label, such as a fluorescein, a phosphor, a rhodamine, or a polymethine dye derivative. Examples of commercially available fluorescent labels include, for example, fluorescent phosphoramidites such as FluorePrime (Amersham Pharmacia, Piscataway, N.J.),
30 Fluoredate (Millipore, Bedford, Mass.), FAM (ABI, Foster City, Calif.), and Cy3 or Cy5 (Amersham Pharmacia, Piscataway, N.J.). In another embodiment, the detectable label is a radiolabeled nucleotide.

In a further preferred embodiment, target polynucleotide molecules from a patient sample are labeled differentially from target polynucleotide molecules of a standard. The standard can comprise target polynucleotide molecules from normal individuals (*i.e.*, those not afflicted with breast cancer). In a highly preferred embodiment, the standard comprises
5 target polynucleotide molecules pooled from samples from normal individuals or tumor samples from individuals having sporadic-type breast tumors. In another embodiment, the target polynucleotide molecules are derived from the same individual, but are taken at different time points, and thus indicate the efficacy of a treatment by a change in expression of the markers, or lack thereof, during and after the course of treatment (*i.e.*, chemotherapy,
10 radiation therapy or cryotherapy), wherein a change in the expression of the markers from a poor prognosis pattern to a good prognosis pattern indicates that the treatment is efficacious. In this embodiment, different timepoints are differentially labeled.

5.3.2.5. HYBRIDIZATION TO MICROARRAYS

Nucleic acid hybridization and wash conditions are chosen so that the target
15 polynucleotide molecules specifically bind or specifically hybridize to the complementary polynucleotide sequences of the array, preferably to a specific array site, wherein its complementary DNA is located.

Arrays containing double-stranded probe DNA situated thereon are preferably subjected to denaturing conditions to render the DNA single-stranded prior to contacting with
20 the target polynucleotide molecules. Arrays containing single-stranded probe DNA (*e.g.*, synthetic oligodeoxyribonucleic acids) may need to be denatured prior to contacting with the target polynucleotide molecules, *e.g.*, to remove hairpins or dimers which form due to self complementary sequences.

Optimal hybridization conditions will depend on the length (*e.g.*, oligomer versus
25 polynucleotide greater than 200 bases) and type (*e.g.*, RNA, or DNA) of probe and target nucleic acids. One of skill in the art will appreciate that as the oligonucleotides become shorter, it may become necessary to adjust their length to achieve a relatively uniform melting temperature for satisfactory hybridization results. General parameters for specific (*i.e.*, stringent) hybridization conditions for nucleic acids are described in Sambrook *et al.*,
30 MOLECULAR CLONING - A LABORATORY MANUAL (2ND ED.), Vols. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989), and in Ausubel *et al.*, CURRENT

PROTOCOLS IN MOLECULAR BIOLOGY, vol. 2, Current Protocols Publishing, New York (1994). Typical hybridization conditions for the cDNA microarrays of Schena *et al.* are hybridization in 5 X SSC plus 0.2% SDS at 65°C for four hours, followed by washes at 25°C in low stringency wash buffer (1 X SSC plus 0.2% SDS), followed by 10 minutes at 25°C in higher stringency wash buffer (0.1 X SSC plus 0.2% SDS) (Schena *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 93:10614 (1993)). Useful hybridization conditions are also provided in, *e.g.*, Tijessen, 1993, HYBRIDIZATION WITH NUCLEIC ACID PROBES, Elsevier Science Publishers B.V.; and Kricka, 1992, NONISOTOPIC DNA PROBE TECHNIQUES, Academic Press, San Diego, CA.

10 Particularly preferred hybridization conditions include hybridization at a temperature at or near the mean melting temperature of the probes (*e.g.*, within 51°C, more preferably within 21°C) in 1 M NaCl, 50 mM MES buffer (pH 6.5), 0.5% sodium sarcosine and 30% formamide.

5.3.2.6. SIGNAL DETECTION AND DATA ANALYSIS

15 When fluorescently labeled probes are used, the fluorescence emissions at each site of a microarray may be, preferably, detected by scanning confocal laser microscopy. In one embodiment, a separate scan, using the appropriate excitation line, is carried out for each of the two fluorophores used. Alternatively, a laser may be used that allows simultaneous specimen illumination at wavelengths specific to the two fluorophores and emissions from the two fluorophores can be analyzed simultaneously (*see* Shalon *et al.*, 1996, "A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization," *Genome Research* 6:639-645, which is incorporated by reference in its entirety for all purposes). In a preferred embodiment, the arrays are scanned with a laser fluorescent scanner with a computer controlled X-Y stage and a microscope objective.

20 Sequential excitation of the two fluorophores is achieved with a multi-line, mixed gas laser and the emitted light is split by wavelength and detected with two photomultiplier tubes. Fluorescence laser scanning devices are described in Schena *et al.*, *Genome Res.* 6:639-645 (1996), and in other references cited herein. Alternatively, the fiber-optic bundle described by Ferguson *et al.*, *Nature Biotech.* 14:1681-1684 (1996), may be used to monitor mRNA

25 abundance levels at a large number of sites simultaneously.

30

Signals are recorded and, in a preferred embodiment, analyzed by computer, *e.g.*, using a 12 or 16 bit analog to digital board. In one embodiment the scanned image is despeckled using a graphics program (*e.g.*, Hijaak Graphics Suite) and then analyzed using an image gridding program that creates a spreadsheet of the average hybridization at each wavelength at each site. If necessary, an experimentally determined correction for “cross talk” (or overlap) between the channels for the two fluors may be made. For any particular hybridization site on the transcript array, a ratio of the emission of the two fluorophores can be calculated. The ratio is independent of the absolute expression level of the cognate gene, but is useful for genes whose expression is significantly modulated in association with the different breast cancer-related condition.

5.4. THERAPEUTIC REGIMENS SPECIFIC TO PATIENT SUBSETS

The benefit of identifying subsets of individuals that have a common condition, followed by identification of sets of genes informative for those particular subsets of individuals, is that such subdivision and identification tends to more accurately identify the subset of genes responsible for, or most closely associated with, a particular form of the condition. For example, breast cancer is a complex condition brought about by several different molecular mechanisms. ER⁺ individuals, particularly ER⁺, ER/AGE high individuals, show an increased level of expression of cell cycle-control genes, and the expression of these genes is highly informative for prognosis in this patient subset (*see* Examples). In ER⁻ individuals, however, the expression of these genes is not informative for prognosis.

The set of informative markers, therefore, can be used to assign a particular course of therapy to an individual, *e.g.*, an individual having breast cancer, depending upon the condition subset into which the individual is classified. In one embodiment, therefore, the invention provides a method of assigning a course of therapy to an individual having a condition, said method comprising classifying the individual into one of a plurality of subsets of a condition, wherein a plurality of informative genes has been identified for at least one of said subsets; and assigning a course of therapy known or suspected to be effective for treating the subset of the condition associated with those genes. In a specific embodiment, said condition is breast cancer, said patient subset is ER⁺, ER/AGE high status, and said course of therapy comprises the administration of one or more compounds known or suspected to be

effective at arresting the cell cycle. In a more specific embodiment, said one or more compounds comprises taxol or a vinca alkaloid.

Of course, any course of therapy selected or assigned on the basis of the above phenotypes and gene expression may be supplemented by other treatments or courses of therapy relevant to or known or suspected to be effective in the treatment of the condition. For example, the treatment of breast cancer may additionally comprise surgery, either tissue-preserving or radical, radiation treatment, chemotherapy other than that suggested by gene expression analysis, or any other therapy or treatment known or suspected to be effective.

5.5. CLINICAL TRIALS AND EPIDEMIOLOGICAL STUDIES

The method of the present invention may also be used to assign individuals to categories within a clinical trial, epidemiological study or the like. For example, individuals may be distinguished according to a characteristic of a condition, such as the presence or absence of specific proteins (*e.g.*, estrogen receptor) or tissue structures (*e.g.*, lymph nodes), and with prognosis, and the results of the trial correlated with prognosis. In a specific example, the condition is breast cancer, the characteristic is the presence of the estrogen receptor, and the prognosis is the expected occurrence or non-occurrence of metastases within a given period, for example, five years, after initial diagnosis. In another specific example, the condition is obesity, the characteristics are 24-hour energy expenditure, and the prognosis is the expected occurrence of heart disease or diabetes. In another specific example, the condition is a neurodegenerative disease, the characteristic is exposure to a particular range of concentration of an environmental toxin, and the prognosis is expected occurrence or degree of loss of motor function. In each case, the characteristics and expected outcome are used to assign the individual to a category within a clinical trial or epidemiological study.

Thus, the invention provides a method for assigning an individual to one of a plurality of categories in a clinical trial, comprising classifying the individual into one of a plurality of condition categories differentiated by at least one genotypic or phenotypic characteristic of the condition; determining the level of expression, in a sample derived from said individual, of a plurality of genes informative for said condition category; determining whether said level of expression of said plurality of genes indicates that the individual has a good prognosis or a poor prognosis; and assigning the individual to a category in a clinical trial on the basis of prognosis.

In a specific embodiment, the invention provides a method of assigning an individual to a category in a breast cancer clinical trial, said method comprising: (a) classifying said individual as ER⁻, *BRCA1*, ER⁻, sporadic; ER⁺, ER/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻; (b) determining for said individual the level of expression of at least two genes for which markers are listed in Table 1 if said individual is classified as ER⁻, *BRCA1*; Table 2 if said individual is classified as ER⁻, sporadic; Table 3 if said individual is classified as ER⁺, ER/AGE high; Table 4 if said individual is classified as ER⁺, ER/AGE low, LN⁺; or Table 5 if said individual is classified as ER⁺, ER/AGE low, LN⁻; (c) determining whether said individual has a pattern of expression of said at least two genes that correlates with a good prognosis or a poor prognosis; and (d) assigning said individual to at least one category in a clinical trial if said individual has a good prognosis, and assigning said individual to a second category in said clinical trial if said individual has a poor prognosis. In a more specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of the classification of said individual as determined in step (a). In another more specific embodiment, said individual is additionally assigned to a category in said clinical trial on the basis of any other clinical, phenotypic or genotypic characteristic of breast cancer. In another more specific embodiment, the method additionally comprises determining in said cell sample the level of expression, relative to a control, of a second plurality of genes for which markers are not found in Tables 1-5, wherein said second plurality of genes is informative for prognosis of breast cancer, and determining from the expression of said second plurality of genes, in addition to said first plurality of genes, whether said individual has a good prognosis or a poor prognosis.

5.6. KITS

The present invention further provides for kits comprising the marker sets above. The components of the kits of the present invention are preferably contained in sealed containers. In a preferred embodiment, the kit comprises a microarray ready for hybridization to target polynucleotide molecules. In specific embodiments, the kit may comprise any of the microarrays described in detail in Section 5.3.2. Where proteins are the target molecules, the kit preferably comprises a plurality of antibodies for binding to specific condition-related proteins, and means for identifying such binding (*e.g.*, means for performing a sandwich assay, ELISA, RIA, *etc.*). Such antibodies may be provided, for example, individually or as part of an antibody array. The kit may additionally comprise software for the data analyses described above, as described in detail in Section 5.7. The kit preferably contains one or

more controls. Such a control may be an artificial population of marker-related or marker-derived polynucleotides suitable for hybridization to a microarray, wherein the markers are related to or relevant to the condition of interest (for example, breast cancer). The control may also, or alternatively, be a set of expression values stored on a computer disk or other
5 storage medium.

The kits of the invention may be primarily diagnostic in nature; that is, they may assist a physician or researcher in determining a characteristic, for example, the prognosis, of a condition of interest, the likely response to a therapeutic regimen, the likely outcome of exposure to an environmental condition, such as toxin exposure, etc. The kits of the
10 invention may also be used to classify individuals, for example, to place individuals into different groups in a clinical trial. The use of each kit is determined by the markers, microarrays, controls, etc. included.

5.7. COMPUTER-FACILITATED ANALYSIS

The analytic methods described in the previous sections can be implemented by use of
15 the following computer systems and according to the following programs and methods. A computer system comprises internal components linked to external components. The internal components of a typical computer system include a processor element interconnected with a main memory. For example, the computer system can be based on an Intel 8086-, 80386-, 80486-, Pentium™, or Pentium™-based processor with preferably 32 MB or more of main
20 memory. The computer system may also be a Macintosh or a Macintosh-based system, but may also be a minicomputer or mainframe.

The external components preferably include mass storage. This mass storage can be one or more hard disks (which are typically packaged together with the processor and memory). Such hard disks are preferably of 1 GB or greater storage capacity. Other external
25 components include a user interface device, which can be a monitor, together with an inputting device, which can be a "mouse", or other graphic input devices, and/or a keyboard. A printing device can also be attached to the computer.

Typically, a computer system is also linked to network link, which can be part of an Ethernet link to other local computer systems, remote computer systems, or wide area
30 communication networks, such as the Internet. This network link allows the computer system to share data and processing tasks with other computer systems.

Loaded into memory during operation of this system are several software components, which are both standard in the art and special to the instant invention. These software components collectively cause the computer system to function according to the methods of this invention. These software components are typically stored on the mass storage device. A software component comprises the operating system, which is responsible for managing computer system and its network interconnections. This operating system can be, for example, of the Microsoft Windows® family, such as Windows 3.1, Windows 95, Windows 98, Windows 2000, or Windows NT, or may be of the Macintosh OS family, or may be UNIX, a UNIX derivative such as LINUX, or an operating system specific to a minicomputer or mainframe. The software component represents common languages and functions conveniently present on this system to assist programs implementing the methods specific to this invention. Many high or low level computer languages can be used to program the analytic methods of this invention. Instructions can be interpreted during run-time or compiled. Preferred languages include C/C++, FORTRAN and JAVA. Most preferably, the methods of this invention are programmed in mathematical software packages that allow symbolic entry of equations and high-level specification of processing, including some or all of the algorithms to be used, thereby freeing a user of the need to procedurally program individual equations or algorithms. Such packages include Matlab from Mathworks (Natick, MA), Mathematica® from Wolfram Research (Champaign, IL), or S-Plus® from Math Soft (Cambridge, MA). Specifically, the software component includes the analytic methods of the invention as programmed in a procedural language or symbolic package.

The software to be included with the kit comprises the data analysis methods of the invention as disclosed herein. In particular, the software may include mathematical routines for marker discovery, including the calculation of similarity values between clinical categories (*e.g.*, prognosis) and marker expression. The software may also include mathematical routines for calculating the similarity between sample marker expression and template marker expression, using array-generated fluorescence data, to determine the clinical classification of a sample.

Additionally, the software may also include mathematical routines for determining the prognostic outcome, and the recommended therapeutic regimen, for an individual with a condition of interest. In the specific example of breast cancer, the mathematical routines

would determine the prognostic outcome and recommended therapeutic regimen for an individual having breast cancer. Such breast cancer-specific software would include instructions for the computer system's processor to receive data structures that include the level of expression of five or more of the marker genes listed in any of Tables 1-5 in a breast cancer tumor sample obtained from the breast cancer patient; the mean level of expression of the same genes in a control or template; and the breast cancer patient's clinical information, including age, lymph node status and ER status. The software may additionally include mathematical routines for transforming the hybridization data and for calculating the similarity between the expression profile of the marker genes in the patient's breast cancer tumor sample and a template. In a specific embodiment, the software includes mathematical routines for calculating a similarity metric, such as a coefficient of correlation, representing the similarity between the expression profile for the marker genes in the patient's breast cancer tumor sample and the template, and expressing the similarity as that similarity metric.

The software preferably would include decisional routines that integrate the patient's clinical and marker gene expression data, and recommend a course of therapy. In one embodiment, for example, the software causes the processor unit to receive expression data for prognosis-related genes in the patient's tumor sample, calculate a metric of similarity of these expression values to the values for the same genes in a template, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to a prognostic group, and, on the basis of the prognostic group, assign a recommended therapeutic regimen. In a specific example, the software additionally causes the processor unit to receive data structures comprising clinical information about the breast cancer patient. In a more specific example, such clinical information includes the patient's age, estrogen receptor status, and lymph node status.

The software preferably causes the processor unit to receive data structures comprising relevant phenotypic and/or genotypic characteristics of the particular condition of interest, and/or of an individual having that condition, and classifies the individual into a condition subset according to those characteristics. The software then causes the processor to receive values for subset-specific markers, to calculate a metric of similarity of the values associated with those markers (*e.g.*, level, abundance, activity, *etc.*) from the individual to a template, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to a prognostic group, and,

on the basis of the prognostic group, assign a recommended therapeutic regimen. In the specific example of breast cancer and a breast cancer patient, the software, in one embodiment, causes the processor unit to receive data structures comprising the patient's age, estrogen receptor status, and lymph node status, and on the basis of this data, to classify the patient into one of the following patient subsets: ER⁻, sporadic; ER⁻, *BRCAl*; ER⁺, AR/AGE high; ER⁺, ER/AGE low, LN⁺; or ER⁺, ER/AGE low, LN⁻. The software then causes the processor to receive expression values for subset-specific prognosis-informative gene expression in the patient's tumor sample, calculate a metric of similarity of these expression values to the values for the same genes in a patient subset-specific template or template, compare this similarity metric to a pre-selected similarity metric threshold or thresholds that differentiate prognostic groups, assign the patient to the prognostic group, and, on the basis of the prognostic group, assign a recommended therapeutic regimen.

Where the template is an expression template comprising expression values for marker genes within a group of patients, *e.g.*, breast cancer patients, the template can comprise either hybridization data obtained at the same time (*i.e.*, in the same hybridization experiment) as the patient's individual hybridization data, or can be a set of hybridization or marker expression values stores on a computer, or on computer-readable media. If the latter is used, new patient hybridization data for the selected marker genes, obtained from initial or follow-up tumor samples, or suspected tumor samples, can be compared to the stored values for the same genes. However, the software may additionally comprise routines for updating the template data set, *e.g.*, to add information from additional breast cancer patients or to remove existing members of the template data set, and, consequently, for recalculating the average expression level values that comprise the template. In another specific embodiment, said template comprises a set of single-channel mean hybridization intensity values for each of said at least five of said genes, stored on a computer-readable medium.

Clinical data relating to a breast cancer patient, or a patient having another type of condition, and used by the computer program products of the invention, can be contained in a database of clinical data in which information on each patient is maintained in a separate record, which record may contain any information relevant to the patient, the patient's medical history, treatment, prognosis, or participation in a clinical trial or study, including expression profile data generated as part of an initial diagnosis or for tracking the progress of the condition, for example, breast cancer, during treatment.

Thus, one embodiment of the invention provides a computer program product for classifying a breast cancer patient according to prognosis, the computer program product for use in conjunction with a computer having a memory and a processor, the computer program product comprising a computer readable storage medium having a computer program

5 mechanism encoded thereon, wherein said computer program product can be loaded into the one or more memory units of a computer and causes the one or more processor units of the computer to execute the steps of (a) receiving a first data structure comprising said breast cancer patient's age, ER status, LN status and tumor type; (b) classifying said patient as ER⁻, sporadic; ER⁻, *BRCA1*; ER+, ER/AGE high; ER+, ER/AGE low, LN+; or ER+, ER/AGE

10 low, LN⁻; (c) receiving a second data structure comprising the levels of expression of at least two genes in a cell sample taken from said breast cancer patient wherein markers for said at least two genes are listed in Table 1 if said patient is classified as ER⁻, sporadic; Table 2 if said patient is classified as ER⁻, sporadic; Table 3 if said patient is classified as ER+, ER/AGE high; Table 4 if said patient is classified as ER+, ER/AGE low, LN+; or Table 5 if

15 said patient is classified as ER+, ER/AGE high, LN⁻; (d) determining the similarity of the expression profile of said at least two genes to a template expression profile of said at least two genes to obtain a patient similarity value; (e) comparing said patient similarity value to selected first and second threshold values of similarity, respectively, wherein said second similarity threshold indicates greater similarity to said template expression profile than does

20 said first similarity threshold; and (f) classifying said breast cancer patient as having a first prognosis if said patient similarity value exceeds said said second threshold similarity values, a second prognosis if said patient similarity value exceeds said first threshold similarity value but does not exceed said second threshold similarity value, and a third prognosis if said patient similarity value does not exceed said first threshold similarity value. In a specific

25 embodiment of said computer program product, said first threshold value of similarity and said second threshold value of similarity are values stored in said computer. In another more specific embodiment, said first prognosis is a "very good prognosis," said second prognosis is an "intermediate prognosis," and said third prognosis is a "poor prognosis," and wherein said computer program mechanism may be loaded into the memory and further cause said one or

30 more processor units of said computer to execute the step of assigning said breast cancer patient a therapeutic regimen comprising no adjuvant chemotherapy if the patient is lymph node negative and is classified as having a good prognosis or an intermediate prognosis, or comprising chemotherapy if said patient has any other combination of lymph node status and expression profile. In another specific embodiment, said computer program mechanism may

be loaded into the memory and further cause said one or more processor units of the computer to execute the steps of receiving a data structure comprising clinical data specific to said breast cancer patient. In a more specific embodiment, said single-channel hybridization intensity values are log transformed. The computer implementation of the method, however, may use any desired transformation method. In another specific embodiment, the computer program product causes said processing unit to perform said comparing step (e) by calculating the correlation between the expression profile of said genes in said cell sample taken from said breast cancer patient and the template expression profile of the same genes. In another specific embodiment, the computer program product causes said processing unit to perform said comparing step (e) by calculating the distance between the expression profile of said genes in said cell sample taken from said breast cancer patient and the template expression profile of the same genes.

Of course, the above breast cancer-specific examples are not limiting; analogous computer systems, software, and data analysis methods may be utilized for any condition of interest. For example, analogous software may be used to determine the prognosis of any other type of cancer, or of any other non-cancer diseases or conditions, using markers, expression level data and controls specific for that cancer, non-cancer disease or condition.

In an exemplary implementation, to practice the methods of the present invention, a user first loads experimental data into the computer system. These data can be directly entered by the user from a monitor, keyboard, or from other computer systems linked by a network connection, or on removable storage media such as a CD-ROM, floppy disk (not illustrated), tape drive (not illustrated), ZIP[®] drive (not illustrated) or through the network. Next the user causes execution of expression profile analysis software which performs the methods of the present invention.

In another exemplary implementation, a user first loads experimental data and/or databases into the computer system. This data is loaded into the memory from the storage media or from a remote computer, preferably from a dynamic geneset database system, through the network. Next the user causes execution of software that performs the steps of the present invention.

Additionally, because the data obtained and analyzed in the software and computer system products of the invention may be confidential, the software and/or computer system

preferably comprises access controls or access control routines, such as password protection and preferably, particularly if information is to be transmitted between computers, for example, over the Internet, encryption of the data by a suitable encryption algorithm (*e.g.*, PGP).

Alternative computer systems and software for implementing the analytic methods of this invention will be apparent to one of skill in the art and are intended to be comprehended within the accompanying claims. In particular, the accompanying claims are intended to include the alternative program structures for implementing the methods of this invention that will be readily apparent to one of skill in the art.

6. EXAMPLES

6.1. EXAMPLE 1: IDENTIFICATION OF PHENOTYPIC SUBSETS AND INFORMATIVE GENESETS FOR EACH

Materials and Methods

Tumor Samples:

311 cohort samples were collected from breast cancer patients. Selection criteria for sporadic patients (*i.e.*, those not identified as having a *BRCA1*-type tumor; $n = 291$) included: primary invasive breast carcinoma less than 5 cm (T1 or T2); no axillary metastases (N0); age at diagnosis of less than 55 years; calendar year of diagnosis 1983-1996; and no previous malignancies. All patients were treated by modified radical mastectomy or breast-conserving treatment. *See van't Veer et al., Nature 415:530 (2002).* Selection criteria for hereditary (*i.e.*, *BRCA1*-type; $n = 20$) tumors included: carriers of germline mutation in *BRCA1* or *BRCA2*, and primary invasive breast carcinoma. *van't Veer, supra.* Additionally, for development of a classifier for the *BRCA1* group, 14 *BRCA1* samples previously identified (*see van't Veer, supra*) were added to the 20 *BRCA1* type samples to increase sample size. Those 14 samples also satisfy the conditions that they are ER negative and age less than 55 years old.

Data analysis:

Sample sub-grouping: As shown in FIG. 1, tumor samples were first divided into ER+ and ER⁻ branches since this is the dominant gene expression pattern. In the ER⁻ branch, the samples were further divided into "BRCA1 mutation like" and "Sporadic like" categories

using the expression templates and 100 genes previously identified as optimal for determining *BRCA1* status. See van't Veer *et al.*, *Nature* 415:530 (2002). In the ER+ category, samples were divided by ER vs. age distribution (see below) into two groups, "ER/AGE low" and "ER/AGE high." Within the "ER/AGE low" group, samples were
 5 further divided according to the lymph node status into two sub-groups: lymph node negative (0 lymph nodes; LN-) and positive (> 0 lymph nodes; LN+) group.

The result of these divisions was five distinctive sub-groups: "ER⁻, sporadic" ($n = 52$), "ER⁻, BRCA1" ($n = 34$), "ER+, ER/AGE high" ($n = 83$), "ER+, ER/AGE low, LN⁻" ($n = 81$), and "ER+, ER/AGE low, LN+" ($n = 75$). A few samples with a specific ER vs. age distribution
 10 in "ER+, ER/AGE low, LN+" group were further excluded to develop a classifier, see below for details.

Estrogen receptor level: Estrogen receptor gene expression level was measured by a 60mer oligo-nucleotide on a microarray. Since every individual sample was compared to a pool of all samples, the ratio to pool was used to measure the relative level. A threshold of -0.65
 15 on $\log_{10}(\text{ratio})$ was used to separate the ER+ group from ER⁻ group. See van't Veer *et al.*, *Nature* 415:530 (2002).

Grouping by ER vs. age distribution: Samples were not uniformly distributed in ER vs. age space among the ER+ samples (FIG. 2). First, it appeared that the ER level increases with age, as there were few samples from young individuals having a high ER expression level.
 20 For example, in the 35 to 40 years age group, samples having a $\log(\text{ratio})$ of ER > 0.2 are relatively few as compared to the 40 to 45 age group. In the set of samples used, the $40 < \text{age} \leq 45$ group contains 30 samples having $\log(\text{ratio})$ ER values between -0.2 to 0.2 , and 28 samples having values greater than 0.2 , whereas the $35 < \text{age} \leq 40$ group includes 24 samples with values between -0.2 to 0.2 , but only 6 samples with values of greater than 0.2 (Fisher's
 25 exact test P-value: 1%). The increasing ER level with age may simply due to the fact that estrogen levels decrease with age, and the estrogen receptor level rises in compensation.

There also appear to be at least two groups of patients, as indicated by the solid line separating the two in FIG. 2A. A bimodality test of the separation indicated by the solid line yielded P-value < 10^{-4} . Each of these two groups has its own trend between the ER level and
 30 age. The solid line can be approximated by $\text{ER} = 0.1(\text{age} - 42.5)$. Patients having values

above the solid line are referred to as the “ER/AGE high” group, and the patients below the line as the “ER/AGE low” group.

Prognosis in each group:

Feature selection and performance evaluation: For the prognosis in each group, non-

- 5 informative genes were filtered in each group of patients. Specifically, only genes with $|\log_{10}(\text{ratio})| > \log_{10}(2)$ and P-value (for $\log(\text{ratio}) \neq 0$) < 0.01 in more than 3 experiments were kept. This step removed all genes that never had any significant change across all samples. The second step used a leave-one-out cross validation (LOOCV) procedure to optimize the number of reporter genes (features) in the classifier and to estimate the
- 10 performance of the classifier in each group. The feature selection was included inside the loop of each LOOCV process. The final “optimal” reporter genes were selected using all of the “training samples” as the result of “re-substitution” because one classifier was needed for each group.

- Selection of training samples: Only the samples from patients who had metastases within 5
- 15 years of initial diagnosis (3 years for “ER⁻, sporadic” samples; *i.e.*, the “poor outcome” group), or who were metastases-free with more than 5 years of follow-up time (*i.e.*, the “good outcome” group), were used as the training set. Because the average expression levels for informative genes among patients who were metastasis-free, or who had early metastases, were used as expression templates for prediction, the training samples for the ER+ samples
- 20 were further limited to those samples that could also be correctly classified by the first round of LOOCV process. For the “ER⁻, sporadic” samples, no such iteration was done because no improvement was observed. For the “ER⁻, BRCA1” samples, an iteration was done, but the training samples in the second iteration were limited to the correctly predicted good outcome samples from the first round of LOOCV, and all the poor outcome samples with
- 25 metastases time less than 5 years. Further limitation of the poor outcome samples was not performed because of the small number of poor samples and the absence of improvement by such limitation. In the first round of LOOCV, except for the “ER⁻, sporadic” group, the number of features was fixed at 50 genes. A patient was predicted to have a favorable outcome, that is, no metastases within five years of initial diagnosis, if the expression of the
- 30 reporter genes in a sample from the individual was more similar to the “average good profile” than the “average poor profile”, and a poor outcome, that is, a metastasis within five years, if

the expression of the reporter genes in the sample was more similar to the “average poor profile” than the “average good profile”.

The justification for such an iteration operation is threefold. First, biologically, there are always a few individuals with specific reasons (different from the vast majority) to stay
5 metastases free or to develop metastases. Second, statistically, most groups of patients include outliers that don’t follow the distribution of the majority of samples. Third, methodologically, the iteration operation is very similar to the idea of “boosting”, but instead of increasing the weights of the samples predicted wrong, emphasis is placed on the well behaved samples for selecting features and training the classifier. Since this process was
10 used to select “training samples”, and the performance was evaluated using the LOOCV (including the feature selection) after the training sample being fixed, there is no issue of over-fitting involved in our procedures. This method of iteration is thus more likely to reveal the dominant mode to metastases within each group.

Error rate and odds ratio, threshold in the final LOOCV: Unless otherwise stated, the error
15 rate is the average error rate from two populations: (1) the number of poor outcome samples misclassified as good outcome samples, divided by the total number of poor outcome samples; and (2) the total number of good outcome samples misclassified as poor outcome samples, divided by the total number of good samples. Two odds ratios are reported for a given threshold: (1) the overall odds ratio and (2) the 5 year odds ratio. The 5 year odds
20 ratio was calculated from samples from individuals that were metastases free for more than five years, and who experienced metastasis within 5 years. The threshold was applied to **cor1 – cor2**, where “cor1” stands for the correlation to the “average good profile” in the training set, and “cor2” stands for the correlation to the “average poor profile” in the training set.

25 The threshold in the final round of LOOCV was defined using the following steps:
(1) For each of the N sample *i* left out for training, features based on the training set were selected. (2) Given a feature set, an incomplete LOOCV with N-1 samples was performed (only the “average poor profile” and “average good profile” is varied depending on whether the left out sample is in the training set or not). (3) The threshold based on the minimum
30 error rate from N-1 samples was determined, and that threshold was assigned to sample *i* in step (1). (4) The median threshold from all N samples was taken, and designated the final threshold. FIGS. 3-7 present detailed information about classifiers for the 5 groups: “ER⁻,”

sporadic", "ER⁻, BRCA1", "ER⁺, ER/age high", "ER⁺, ER/age low, LN⁻", "ER⁺, ER/age low, LN⁺". Tables 1-5 (*see* Section 5.3) list the final optimal reporter genes for each of the 5 classifiers for each of the five patient subsets. Table 6, below, summarizes the performance of each of the five classifiers together with thresholds used in each classifier.

5 Table 6. Performance of classifiers for each patient subset.

Classifier	Optimal # of Genes	(C1-C2) Threshold	Metastasis Free	# of Samples	TP	FP	FN	TN	Odds Ratio	95% C.I.
ER ⁺ , ER/AGE high	50	1.22	Overall	83	31	14	5	33	14.61	4.71-45.36
			5 year	71	24	11	3	33	24.00	6.03-95.46
ER ⁺ , ER/AGE low, LN ⁻	65	0.38	Overall	81	14	6	6	55	21.39	5.98-76.52
			5 year	73	11	4	5	53	29.15	6.73-126.33
ER ⁺ , ER/AGE low, LN ⁺	50	-0.12	Overall	56	7	4	6	39	11.38	2.54-50.94
			5 year	48	5	4	3	36	15.00	2.57-87.64
ER ⁻ , sporadic	20	-0.01	Overall	52	18	7	7	29	7.35	2.16-25.04
			5 year	45	16	5	6	18	9.60	2.45-37.58
ER ⁻ , BRCA1	10	-0.37	Overall	34	6	3	3	22	14.67	2.34-92.11
			5 year	22	6	1	3	12	24.00	2.04-282.68

TP: True positive

FP: False positive

FN: False negative

TN: True negative

- 10 Classification method: All classifiers described herein, feature selection and optimization were included inside the LOOCV loop. Classifier performance was based on the LOOCV results. The profile based on the selected features from each patient was compared to the "average good profile" and "average poor profile" (by correlation) to determine its predicted outcome.

- 15 Correlation calculation: The correlation between each gene's expression log(ratio) and the endpoint data (final outcome) was calculated using the Pearson's correlation coefficient. The correlation between each patient's profile and the "average good profile" and "average poor profile" is the cosine product (no mean subtraction).

Results:

We employed the comprehensive prognosis strategy on microarray expression profiles of 311 patients diagnosed before age 55 that were all part of previous studies establishing and validating a 70-gene prognosis profile. *See van 't Veer et al., Nature* 415:530 (2002); *van de Vijver et al., N. Engl. J. Med.* 347:1999 (2002). In addition, 14
5 known *BRCA1* samples from the *Nature* study were included in defining the prognosis classifier for the *BRCA1* group. The overview of the stratifications is shown in FIG. 1. In each of the patient subsets, prognosis classifiers were developed and performance was evaluated by leave-one-out cross-validation. The biological make up of each of the classifiers was also examined.

10 During the process to decide whether a particular clinical parameter should be used for the next stratification, our objectives were twofold: (1) identification of homogeneous prognosis patterns; and/or (2) improved prognosis in the subsets. There is a subtle balance between these two objectives because smaller groups will likely lead to uniform patterns within the group but have increasingly limited predictive power. With the exception of the *BRCA1*
15 subset, each group in our stratification contained 50 or more samples.

The first layer of stratification was based on the estrogen receptor level. We and others previously observed that estrogen receptor expression has a dominant effect on overall gene expression in breast cancer as seen in hierarchical clustering. *van 't Veer et al., Nature* 415:530 (2002); *Perou et al., Nature* 406:747 (2000); *Gruvberger et al., Cancer Res.* 61:5979
20 (2001). In our previous analysis up to 2500 genes are significantly correlated with ER expression levels in tumor. *van 't Veer et al., Nature* 415:530 (2002). According to the threshold defined previously (*van de Vijver et al., N. Engl. J. Med.* 347:1999 (2002)), samples were first divided into two groups according to the estrogen receptor level as measured by the oligo probe (accession number: NM_000125) on the array; samples with
25 $\log(\text{ratio}) > -0.65$ belong to the ER⁺ group, and the rest belong to ER⁻ group). This resulted in 239 samples in the ER⁺ group and 72 samples in the ER⁻ group.

In the ER⁺ branch we observed that when displaying ER expression level as a function of age, at least two subgroups appear to exist. (In general, any bimodality in the clinical data is useful.) We therefore decided to stratify the tumors according this bimodality
30 (*see* FIG. 2). The group of ER⁺ patients having a high ER/AGE ratio was designated the "ER/AGE high" group (83 samples), and the remaining group of patients was designated "ER/AGE low" group (156 samples).

Within the “ER/age high” group, we identified a group of prognosis reporter genes that highly correlated with the outcome (*see* Table 3). Moreover, the expression of these genes appeared to be very homogeneous, as indicated by high similarity in expression among those genes. *See* FIG 2A. Leave-one-out cross validation including reporter selection
5 yielded an odds ratio of 14.6 (95%CI: 4.7-45.4) and 5 year odds ratio of 24.0 (95%CI: 6.0-95.5). Examination of those reporter genes reveals they are mostly the cell cycle genes which are highly expressed in the poor outcome tumors. It is worth noting that even though this group includes LN+ and LN- individuals, and mixed treatment, the incidence of distant metastases is predicted by a biologically uniform set of genes, possibly indicating that
10 proliferation is the prime driving force for disease progression. Also even though variation in these genes is observed in other tumor subgroups this is generally not correlated with outcome in those settings (*see* below).

In the “ER/age low” group, no predictive pattern was found in the whole group; thus, the samples were further stratified into LN- (81 samples, referred to as “ER/age low LN⁻”) and LN+ (75 samples, referred to as “ER/age low LN⁺”) group.
15

Within the “ER/age low LN⁻” group, a group of genes was identified that was uniformly co-regulated, and which correlated with the outcome. Leave-one-out cross-validation (including feature selection) yielded an odds ratio of 21.4 (95% CI: 6.0-76.5) and 5 year odds ratio of 29.2 (95% CI: 6.7-126.3). This group of genes is also enriched for
20 individual biological functions (*see* below).

For the “ER/age low LN⁺” subset, an informative set of genes (*see* Table 4) was obtained after exclusion of several samples from older individuals having low ER levels. These samples are indicated in FIG. 2A as those lying below the dashed line (approximated as $ER < 0.1 * (age - 50)$). 56 samples remained after the exclusion. This sample set allowed the
25 identification of a group of genes with a highly homogeneous pattern that is useful for prognosis (overall odds ratio: 11.4 (2.5-50.9), 5 year odds ratio: 15.0 (2.6-87.6)). This suggests again that ER vs. age is an important combination for stratifying breast cancer patients. The reporter genes involved in this classifier also correlated with the clinical measure of the degree of lymphocytic infiltration (data not shown). The prediction in this
30 group is not as strong as other positive groups, which may indicate the primary tumor carries weaker information about the metastases for this group of patients, and the metastases may be started from or influenced by tumors already in lymph nodes.

In the ER⁻ branch, because a portion of the samples are “BRCA1-like,” it is natural to divide the samples into “BRCA1-like” and “sporadic like”. To perform the classification, the BRCA1/sporadic tumor type classifier described in Roberts *et al.*, “Diagnosis and Prognosis of Breast Cancer Patients,” International Publication No. WO 02/103320, which is hereby
5 incorporated by reference in its entirety, to segregate the ER⁻ cohort samples. 52 out of the 72 ER⁻ samples were found to be “sporadic like” and 20 were found to be “BRCA1-like”. Interestingly, the “sporadic like” group is enriched for erbb2 mutations (data not shown).

Within the “ER⁻, sporadic” group, no homogeneous prognosis pattern was identified; however, 20 genes were identified that are highly predictive of the tumor outcome (*see* Table
10 2). Leave-one-out cross-validation including feature selection yielded an odds ratio of 7.4 (95% CI 2.2-25.0) and 5 year odds ratio 9.6 (2.5 – 37.6). This result represents a significant improvement in prognosis compared to the previously-identified 70 gene prognosis classifier (*see* Roberts *et al.*, International Publication No. WO 02/103320; van 't Veer *et al.*, *Nature* 415:530 (2002)) which has no within-group prognostic power for the ER⁻ patient subset.
15 The fact that 20 genes predict outcome and that there is no homogeneous (and apparent biological) pattern in this group probably indicates multiple mechanisms of metastasis in this group. Gene annotation indicates that genes included may be involved in invasion, energy metabolism and other functions.

For the “ER⁻, BRCA1-like” group, we added 14 BRCA1 mutation carrier samples
20 from our previous study to increase the number of samples. Those 14 extra samples also satisfy our selection criteria: ER negative and age less than 55 years. The leave-one-out cross validation process identified 10 genes that are predictive of final outcomes. The overall odds ratio is 14.7 (95% CI: 2.3-92.1) and the 5 year odds ratio is 24.0 (95% CI: 2.0-282.7).

Because no homogeneous gene expression patterns were found in ER⁻ branch, the
25 predictive power of those genes was further validated. One means of further validation was to review the different classifier gene sets for biological interpretations and to identify genes within each classifier that gave indications as to the origins of the tumors.

The “ER⁺, ER/AGE high” group yielded a classifier highly enriched for cell cycle genes with both G1/S and G2/M phases represented. In this group, over-expression of 46 of
30 the 50 genes is associated with disease progression including all the known cell cycle genes. This is consistent with rapid growth being the determinant of metastatic potential. Four

genes in this classifier are anti-correlated with outcome and cell cycle. One of these genes encodes follistatin, which binds to and inhibits activin and other members of the TGF β family (Lin *et al.*, *Reproduction* 126:133 (2003)), the members of which have many functions, including growth stimulation. Tumor grade also accurately predicts metastatic potential in this group (overall odds ratio: 5.9, 95% CI: 2.0-18.0, 5 year odds ratio: 12.5, 95% CI: 2.6-59.3) and is also correlated with the expression level of these genes, which is consistent with rate of growth being the primary determinant of disease progression. This set of genes has a significantly lower correlation with outcome in the other patient subsets, even though coordinate and similarly variable expression is seen. For example, many tumors in the "ER⁻, sporadic" group have high cell cycle and low FST expression, but the expression of these genes in these groups is minimally correlated with outcome, indicating that growth is not the primary determinant of outcome here (*see* FIGS. 8A and 8B).

The ER⁺, ER/AGE low, LN⁻ group yielded a classifier rich in both genes for glycolytic enzymes (12 of 56) and genes induced by hypoxia and/or angiogenesis (14 of 56) with 5 genes falling into both categories. These genes are positively correlated with poor outcome, implying that energy metabolism (glycolysis), angiogenesis and adaptation to hypoxia are critical pathways in this subgroup of tumors. None of these genes appear in the classifiers for the other patient subsets, and there is a much reduced predictive value of these genes in the other tumors, even though coordinate and similarly variable expression is seen (*see* FIG. 8C and 8D).

The implication of the above analyses is that certain well known functions (growth, angiogenesis, energy metabolism) are important in certain tumor types and not in others, and therefore therapies that target these functions will be likely be similarly effective in some tumor subgroups and not in others. For example therapies that target cell cycle progression, such as taxol or the vinca alkaloids, may be optimally effective in the ER⁺, ER/AGE high group, where overexpression of cell cycle genes predominates in the classifier. In contrast, tumor subgroups in which variation in cell cycle expression is not correlated with outcome may be less sensitive to taxol or the vinca alkaloids.

The "comprehensive prognosis" approach significantly improves the prediction error rate when compared with 70 gene classifier (Table 7). To make the comparison fair, we listed two sets of results from the 70 gene classifier. The first results from the use of the same threshold applied to all the patient subsets (threshold previously optimized for false

negative rate); the second one results from the use of a threshold optimized for each patient subset (optimized for average error rate). The comprehensive approach lowered the error rate by at least 6%.

Table 7. Average error rate for the patient subset approach compared with the previously-described 70 gene classifier.

Prognosis method	over all error rate	5 year error rate
70 gene, fix thresh	30.90%	25.70%
70 gene, opt thresh	28.60%	27.60%
Comprehensive	21.50%	19.30%

Fix thresh: use of a fixed threshold in the classifier as previously determined.

Opt threshold: use of a threshold optimized for each sub-group. For the "ER/Age low, LN+" subgroup, 56 samples used for developing the classifier were included here, resulted in 306 samples in total.

6.2. EXAMPLE 2: IDENTIFICATION OF CHEMOTHERAPY RESPONSE SUBPOPULATION IN BREAST CANCER PATIENTS

Predicting responders to a treatment is extremely important in daily clinical therapies, in order to select the right medicine. It is also crucial for pharmaceutical companies running clinical trials on new medicines to target the right population for efficacy. This Example shows that, by combining clinical parameters with gene expression prediction, a subpopulation of patients with characteristics of "ER/age low" and a "poor signature" has an enhanced response to chemotherapy treatment. On the other hand, patients with a poor signature but do not respond to chemotherapy are potential target for clinical trials of new anti-cancer medicines.

Introduction

Gene expression can be used to predict breast cancer outcome. A "70 gene" classifier (see Table 8) that can be used to identify patients with good outcome for avoiding over-treatment of chemotherapy was demonstrated (van 't Veer LJ, Dai H, van de Vijver MJ, et al., Gene expression profiling predicts clinical outcome of breast cancer, Nature 2002; 415: 530-536). This classifier has been verified by a large cohort of samples (van de Vijver MJ, He YD, van't Veer LJ, et al., A gene-expression signature as a predictor of survival in breast

cancer, N Engl J Med 2002; 347: 1999-2009). The result also implies that the treatment effect will be more significant within the “predicted poor” group since patients in the “predicted good” group are already very good, and may not improve further by additional treatment.

5 Recently, it has been demonstrated that ER+ patients can be further divided into subpopulations. The cell proliferation signature has a very strong prognostic power in a group of patients characterized by a high estrogen receptor levels relative to their ages (termed “ER/age high” group), but reduced predictive power in the “ER/age low” group and no predictive power in the ER- patients (U.S. Provisional Patent Application No. 60/550,810
10 by Dai et al., filed March 5, 2004, attorney docket no. 9301-229-888, which is incorporated by reference in its entirety).

Data set

A cohort of 311 breast cancer samples were used (van 't Veer LJ, Dai H, van de Vijver MJ, et al., Gene expression profiling predicts clinical outcome of breast cancer, Nature
15 2002; 415: 530-536; van de Vijver MJ, He YD, van't Veer LJ, et al., A gene-expression signature as a predictor of survival in breast cancer, N Engl J Med 2002; 347: 1999-2009). Those samples were from patients younger than 55 years of age at the time of cancer diagnosis. Among them, 110 patients received chemotherapy.

Results

20 The following section shows that by combining the microarray prediction with the patient stratification based on clinical parameters, a subpopulation of breast cancer patients with enhanced chemotherapy treatment response can be identified.

The treatment response was measured by the improvement in the metastasis free probability at 10 years after the diagnosis, between the treated patients and the untreated
25 patients.

Chemotherapy has limited effect in breast cancer patients less than 55 years of age.

Shown in Figure 9 is the metastasis free probability as a function of time after diagnosis (Kaplan-Meier plot) for the entire set of patients. Patients were divided into two groups according to whether or not they received chemotherapy. Within the entire set of 311

patients younger than 55 years, 110 received chemotherapy. The treatment effect is very small, with the difference in 10 year metastasis free probability being only ~3%.

The “70 gene” prognostic classifier enhances the treatment effect. The purpose of the “70 gene” prognostic classifier (Table 8) was to save patients from unnecessary chemotherapy if they have a “good signature”. In other words, there may not be further benefit to patients in a good prognosis patient group if they potentially have a very good outcome. Hence, treatment should be given to the “predicted poor” group since they could potentially benefit from it. Results in Figure 10 validate the idea. Figure 10a displays the treatment effect of the “predicted good” group. The group was selected by high correlation to the “average expression of good outcome patients” as determined by 70 gene marker set (column C1 of Table 1) with correlation > 0.5. There were 84 patients within this group, and 44 received chemotherapy. The treatment showed no positive effect in reducing the metastasis rate.

However, the patients with “poor signature” demonstrate an improved treatment response (Figure 10b) as compared to Figure 9. There are 226 patients with “poor signature” (correlation < 0.5), and 81 received chemotherapy. The 10 year metastasis free probability for those received chemotherapy is 9% higher than those did not.

In this example, a threshold of 0.5 in correlation was used to select patients with “poor signature”. This threshold is higher than the threshold of 0.4 that was used previously (van 't Veer LJ, Dai H, van de Vijver MJ, et al., Gene expression profiling predicts clinical outcome of breast cancer, Nature 2002; 415: 530-536; van de Vijver MJ, He YD, van't Veer LJ, et al., A gene-expression signature as a predictor of survival in breast cancer, N Engl J Med 2002; 347: 1999-2009), for the purpose of keeping as many patients in the “predicted poor” group as possible for further stratification. The results were very similar if the original threshold of 0.4 was used.

Limiting to ER+ patients gains extra treatment response. Breast cancer patients with positive estrogen receptor status (ER+) have a distinctively different gene expression pattern than those with negative status(ER-). The “70 gene” classifier has a good prognostic power in the ER+ patients, but little prognostic power in the ER- patients (almost all the ER- patients were predicted to have a poor outcome).

The treatment effect is also very different for ER- patients than for ER+ patients with “poor signature”. As shown in Figure 11a, there is no difference in metastasis rate between the treated and untreated patients in the ER- group.

The treatment effect for ER+ patients with a “poor signature” (Figure 11b) is further increased when compared to Figure 10b. There are 155 patients in this group, and 60 received chemotherapy. The metastasis free probability at 10 years for patients with treatment is 14% higher than those without.

“ER/age high” group is ideal for chemotherapy treatment. Recently, it was discovered that the ER+ patients can be further divided into two groups according to their ER levels relative to the age at diagnosis. As shown in Figure 12, patients with relatively high ER level to their age are termed “ER/AGE high” group, and patients with relatively low ER level to their age are termed as “ER/AGE low” group. The “ER/AGE high” group has a relatively poor outcome compared to the “ER/AGE low” group, but the outcome in the “ER/AGE high” group can be accurately predicted by the cell proliferation genes. The same set of genes has a reduced prognostic power in the “ER/AGE low” group.

As shown in Figure 13, chemotherapy has little effect in the patients of “ER/AGE high” group with a “poor signature”. Excluding this group resulted in a further improved effect in the “ER/AGE low” group. In the “ER/AGE low” group, there are 101 patients with “poor signature”, and 39 received chemotherapy. The difference in metastasis free probability at 10 years between the treated and untreated group is 20%.

In summary, by progressively excluding “non-responsive” patient groups as defined by prognosis signature and clinical parameters (ER and age), it is possible to improve the treatment effect (as reported by the increased metastasis free probability at 10 years after diagnosis) from 3% to 20% (Table 9). Roughly, for patients younger than 55 years, there are about 1/3 of patients that do not need a treatment (“good signature”), about 1/3 patients do not respond to the chemotherapy (ER- and “ER/AGE high”) and about 1/3 patients are “responders” that should be treated with chemotherapy (“ER/AGE low”).

Table 9: Treatment effect in various patient groups

	All patients	Patients with "poor signature"	ER+ & "poor signature"	ER+, ER/AGE high & "poor signature"
--	--------------	--------------------------------	------------------------	-------------------------------------

Difference in 10 year metastasis free probability with/without chemo	3%	9%	14%	20%
---	-----------	-----------	------------	------------

7. REFERENCES CITED

All references cited herein are incorporated herein by reference in their entirety and for all purposes to the same extent as if each individual publication or patent or patent
5 application was specifically and individually indicated to be incorporated by reference in its entirety for all purposes.

Many modifications and variations of the present invention can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. The specific embodiments described herein are offered by way of example only, and the invention is to be
10 limited only by the terms of the appended claims along with the full scope of equivalents to which such claims are entitled.

WHAT IS CLAIMED IS:

1. A method for predicting responsiveness of a breast cancer patient to chemotherapy based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii)
5 an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method comprising predicting said patient to exhibit

10 (a) enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is low; or

(b) reduced response to chemotherapy as compared to patients in the general population of breast cancer patients if (i) said ER level is ER⁻, or (ii) said cellular constituent
15 profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is high;

wherein said cellular constituent profile is classified as a poor prognosis profile if said cellular constituent profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template, said good prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular
20 constituents in a plurality of good outcome patients and said poor prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of poor outcome patients, wherein a good outcome patient is a breast cancer patient who has non-occurrence of metastases within a first period of time after initial diagnosis and a poor outcome patient is a patient who has occurrence of
25 metastases within a second period of time after initial diagnosis.

2. The method of claim 1, said method comprising predicting said patient to exhibit enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is low.

3. The method of claim 1, said method comprising predicting said patient to exhibit reduced response to chemotherapy as compared to patients in the general population of breast cancer patients if said ER level is ER⁻.

5 4. The method of claim 1, said method comprising predicting said patient to exhibit reduced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is high.

5. The method of claim 1, further comprising determining said cellular constituent profile, said ER level, and/or, said ER/AGE.

10 6. The method of claim 5, wherein said cellular constituent profile comprises measurements of a plurality of transcripts in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality
15 of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

7. The method of claim 6, wherein said cellular constituent profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a
20 control sample.

8. The method of claim 1, wherein said cellular constituent profile comprises measurements of a plurality of protein species in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and
25 wherein said poor prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of poor outcome patients.

9. The method of claim 7, wherein measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of
30 good outcome patients.

10. The method of claim 9, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a correlation coefficient between said cellular constituent profile and said good prognosis template, wherein said correlation coefficient greater than a correlation threshold indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

11. The method of claim 9, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a distance between said cellular constituent profile and said good prognosis template, wherein said distance less than a given value indicates a high similarity and said distance equal to or greater than said given value indicates a low similarity.

12. The method of claim 10, wherein said correlation threshold is 0.5.

13. The method of claim 12, wherein said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor in said patient relative to expression level of said gene in said control sample, and wherein said ER level is classified as ER⁺ if log₁₀(ratio) of said expression level is greater than -0.65, and wherein said ER level is classified as ER⁻ if log₁₀(ratio) of said expression level is equal to or less than -0.65.

14. The method of claim 13, wherein said gene encoding said estrogen receptor is the estrogen receptor α gene.

15. The method of claim 14, wherein said ER/AGE is classified as high if said ER level as measured by log₁₀(ratio) is greater than $c \cdot (AGE - d)$, and wherein said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (AGE - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

16. The method of claim 15, wherein said estrogen receptor level is measured by an oligonucleotide probe that detects a transcript corresponding to the gene having accession number NM_000125, wherein said control sample is a pool of breast cancer cells of different patients, and wherein $c = 0.1$ and $d = 42.5$.

17. The method of claim 16, wherein said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients.

18. The method of claim 16, wherein said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript of said gene encoding said estrogen receptor.

19. The method of claim 17, wherein said plurality of transcripts consists of
5 transcripts corresponding to at least a portion of the set of genes listed in Table 8.

20. The method of claim 19, wherein said plurality of transcripts consists of all transcripts corresponding to genes listed in Table 8.

21. The method of claim 20, wherein said differential measurements is selected from the group consisting of xdev, log(ratio), error-weighted log(ratio), and mean subtracted
10 log(intensity).

22. The method of claim 21, wherein said first period is 10 years and said second period is 10 years.

23. The method of claim 22, wherein said patient is under the age of 55, and wherein each said good outcome patient and each said poor outcome patient are under the age of 55 at
15 time of diagnosis of breast cancer.

24. A method for selecting a patient for enrollment in a clinical trial of a drug for treating breast cancer based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or
20 estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method comprising selecting a patient for inclusion in said clinical trial if (a) said ER level is ER⁻; or (b) said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high;

25 wherein said cellular constituent profile is classified as a poor prognosis profile if said cellular constituent profile has a low similarity to a good prognosis template or has a high similarity to a poor prognosis template, said good prognosis template comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of good outcome patients and said poor prognosis template

comprising measurements of said plurality of cellular constituents representative of levels of said cellular constituents in a plurality of poor outcome patients, wherein a good outcome patient is a breast cancer patient who has non-occurrence of metastases within a first period of time after initial diagnosis and a poor outcome patient is a breast cancer patient who has
5 occurrence of metastases within a second period of time after initial diagnosis.

25. The method of claim 24, said method comprising selecting a patient for inclusion in said clinical trial if said ER level is ER⁻.

26. The method of claim 24, said method comprising selecting a patient for inclusion in said clinical trial if said cellular constituent profile is a poor prognosis profile, said ER
10 level is ER⁺, and said ER/AGE is high.

27. The method of claim 24, further comprising determining said cellular constituent profile, said ER level, and/or, said ER/AGE.

28. The method of claim 27, wherein said cellular constituent profile comprises measurements of a plurality of transcripts in a sample derived from said patient, wherein said
15 good prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

20 29. The method of claim 28, wherein said cellular constituent profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample.

30. The method of claim 24, wherein said cellular constituent profile comprises
25 measurements of a plurality of protein species in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of poor outcome
30 patients.

31. The method of claim 29, wherein measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of good outcome patients.

32. The method of claim 31, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a correlation coefficient between said cellular constituent profile and said good prognosis template, wherein said correlation coefficient greater than a correlation threshold indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

33. The method of claim 31, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a distance between said cellular constituent profile and said good prognosis template, wherein said distance less than a given value indicates a high similarity and said distance greater than said given value indicates a low similarity.

34. The method of claim 32, wherein said correlation threshold is 0.5.

35. The method of claim 34, wherein said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor in said patient relative to expression level of said gene in said control sample, and wherein said ER level is classified as ER⁺ if $\log_{10}(\text{ratio})$ of said expression level is greater than -0.65, and wherein said ER level is classified as ER⁻ if $\log_{10}(\text{ratio})$ of said expression level is equal to or less than -0.65.

36. The method of claim 35, wherein said gene encoding said estrogen receptor is the estrogen receptor α gene.

37. The method of claim 36, wherein said ER/AGE is classified as high if said ER level is greater than $c \cdot (\text{AGE} - d)$, and wherein said ER/AGE is classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

38. The method of claim 37, wherein said estrogen receptor level is measured by an oligonucleotide probe that detects a transcript corresponding to the gene having accession number NM_000125, wherein said control sample is a pool of breast cancer cells, and wherein $c = 0.1$ and $d = 42.5$.

39. The method of claim 38, wherein said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients.

40. The method of claim 38, wherein said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript corresponding to said gene encoding said estrogen receptor.

41. The method of claim 39, wherein said plurality of transcripts consists of transcripts corresponding to at least a portion of the set of genes listed in Table 8.

42. The method of claim 41, wherein said plurality of genes consists of all transcripts corresponding to genes listed in Table 8.

43. The method of claim 42, wherein said differential measurements is selected from the group consisting of xdev, log(ratio), error-weighted log(ratio), and mean subtracted log(intensity).

44. The method of claim 43, wherein said first period is 10 years and said second period is 10 years.

45. The method of claim 44, wherein said patient is under the age of 55, and wherein each said good outcome patient and each poor outcome patient are under the age of 55 at time of diagnosis of breast cancer.

46. A method for identifying a breast cancer patient as a good or poor candidate for chemotherapy based on one or more of the following (i) a cellular constituent profile comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method comprising

(a) determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, said breast cancer patient is identified as a good candidate for chemotherapy; or

(b) determining whether said ER level is ER⁻, wherein if said ER level is ER⁻, said the breast cancer patient is identified as a poor candidate for chemotherapy; or

(c) determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high, said breast cancer patient is identified as a poor candidate for chemotherapy.

47. The method of claim 46, said method comprising determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is low, said breast cancer patient is identified as a good candidate for chemotherapy.

48. The method of claim 46, said method comprising determining whether said ER level is ER⁻, wherein if said ER level is ER⁻, said the breast cancer patient is identified as a poor candidate for chemotherapy.

49. The method of claims 46, said method comprising determining whether said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high, wherein if said cellular constituent profile is a poor prognosis profile, said ER level is ER⁺, and said ER/AGE is high, said breast cancer patient is identified as a poor candidate for chemotherapy.

50. The method of claim 46, further comprising determining said cellular constituent profile, said ER level, and/or, said ER/AGE.

51. The method of claim 50, wherein said cellular constituent profile comprises measurements of a plurality of transcripts in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of transcripts representative of expression levels of said transcripts in said plurality of poor outcome patients.

52. The method of claim 51, wherein said cellular constituent profile is a differential expression profile comprising differential measurements of said plurality of transcripts in said sample derived from said patient versus measurements of said plurality of transcripts in a control sample.

5 53. The method of claim 50, wherein said cellular constituent profile comprises measurements of a plurality of protein species in a sample derived from said patient, wherein said good prognosis template comprises measurements of said plurality of protein species representative of levels of said protein species in said plurality of good outcome patients, and wherein said poor prognosis template comprises measurements of said plurality of protein
10 species representative of levels of said protein species in said plurality of poor outcome patients.

54. The method of claim 52, wherein measurement of each said transcript in said good prognosis template is an average of expression levels of said transcript in said plurality of good outcome patients.

15 55. The method of claim 54, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a correlation coefficient between said cellular constituent profile and said good prognosis template, wherein said correlation coefficient greater than a correlation threshold indicates a high similarity and said correlation coefficient equal to or less than said correlation threshold indicates a low similarity.

20 56. The method of claim 55, wherein similarity of said cellular constituent profile to said good prognosis template is represented by a distance between said cellular constituent profile and said good prognosis template, wherein said distance less than a given value indicates a high similarity and said distance greater than said given value indicates a low similarity.

25 57. The method of claim 56, wherein said correlation threshold is 0.5.

58. The method of claim 57, wherein said ER level is determined by measuring an expression level of a gene encoding said estrogen receptor in said patient relative to expression level of said gene in said control sample, and wherein said ER level is classified as ER⁺ if log₁₀(ratio) of said expression level is greater than -0.65, and wherein said ER level
30 is classified as ER⁻ if log₁₀(ratio) of said expression level is equal to or less than -0.65.

59. The method of claim 58, wherein said gene encoding said estrogen receptor is the estrogen receptor α gene.

60. The method of claim 59, wherein said ER/AGE is classified as high if said ER level as measured by $\log_{10}(\text{ratio})$ is greater than $c \cdot (\text{AGE} - d)$, and wherein said ER/AGE is
5 classified as low if said ER level is equal to or less than $c \cdot (\text{AGE} - d)$, wherein c is a coefficient, AGE is the age of said patient, and d is an age threshold.

61. The method of claim 60, wherein said estrogen receptor level is measured by an oligonucleotide probe that detects a transcript corresponding to the gene having accession number NM_000125, wherein said control sample is a pool of breast cancer cells, and
10 wherein $c = 0.1$ and $d = 42.5$.

62. The method of claim 61, wherein said control sample is generated by pooling together cDNAs of said plurality of transcripts from a plurality of breast cancer patients.

63. The method of claim 61, wherein said control sample is generated by pooling together synthesized cDNAs of said plurality of transcripts and said transcript corresponding
15 to said gene encoding said estrogen receptor.

64. The method of claim 61, wherein said plurality of transcripts consists of transcripts corresponding to at least a portion of the set of genes listed in Table 8.

65. The method of claim 64, wherein said plurality of genes consists of all transcripts corresponding to genes listed in Table 8.

20 66. The method of claim 65, wherein said differential measurements is selected from the group consisting of $xdev$, $\log(\text{ratio})$, error-weighted $\log(\text{ratio})$, and mean subtracted $\log(\text{intensity})$.

67. The method of claim 66, wherein said first period is 10 years and said second period is 10 years.

25 68. The method of claim 67, wherein said patient is under the age of 55, and wherein each said good outcome patient and each poor outcome patient are under the age of 55 at time of diagnosis of breast cancer.

69. The method of any one of claims 1-68, wherein said chemotherapy is carried out using the CMF combination comprising cyclophosphamide, methotrexate, and 5-fluorouracil.

70. A computer system comprising

a processor, and

5 a memory coupled to said processor and encoding one or more programs,

wherein said one or more programs cause the processor to carry out the method of any one of claims 1-68.

71. A computer program product for use in conjunction with a computer having a processor and a memory connected to the processor, said computer program product
10 comprising a computer readable storage medium having a computer program mechanism encoded thereon, wherein said computer program mechanism may be loaded into the memory of said computer and cause said computer to carry out the method of any one of claims 1-68.

72. A method for predicting responsiveness of a breast cancer patient to chemotherapy based on one or more of the following (i) a cellular constituent profile
15 comprising measurements of a plurality of cellular constituents in a sample derived from said patient, (ii) an ER level, said ER level being a level of an estrogen receptor transcript or estrogen receptor protein in a sample derived from said patient, wherein ER⁺ designates a high ER level and ER⁻ designates a low ER level, and (iii) an ER/AGE, said ER/AGE being a metric of said ER level relative to the age of said patient, said method comprising predicting
20 said patient to exhibit

(a) enhanced response to chemotherapy as compared to patients in the general population of breast cancer patients if said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is low; or

(b) reduced response to chemotherapy as compared to patients in the general
25 population of breast cancer patients if (i) said ER level is ER⁻, or (ii) said cellular constituent profile is a poor prognosis profile, and said ER level is ER⁺, and said ER/AGE is high.
wherein said cellular constituent profile is classified as a good prognosis profile if said cellular constituent profile predicts non-occurrence of metastases in said breast cancer patient within a predetermined period of time after initial diagnosis, and wherein said cellular

constituent profile is classified as a poor prognosis profile if said cellular constituent profile predicts occurrence of metastasis within said predetermined period of time.

1/15

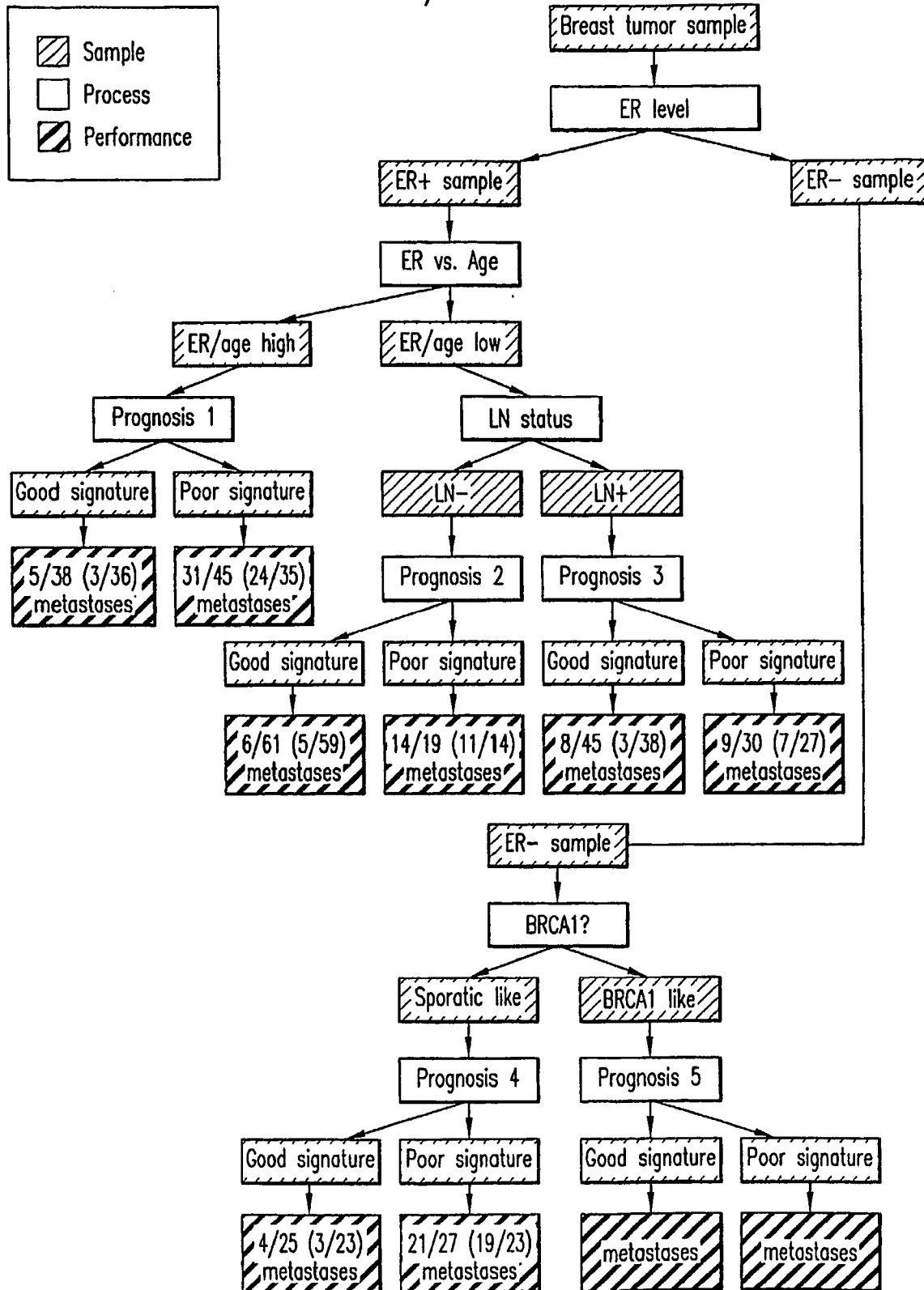


FIG. 1

2/15

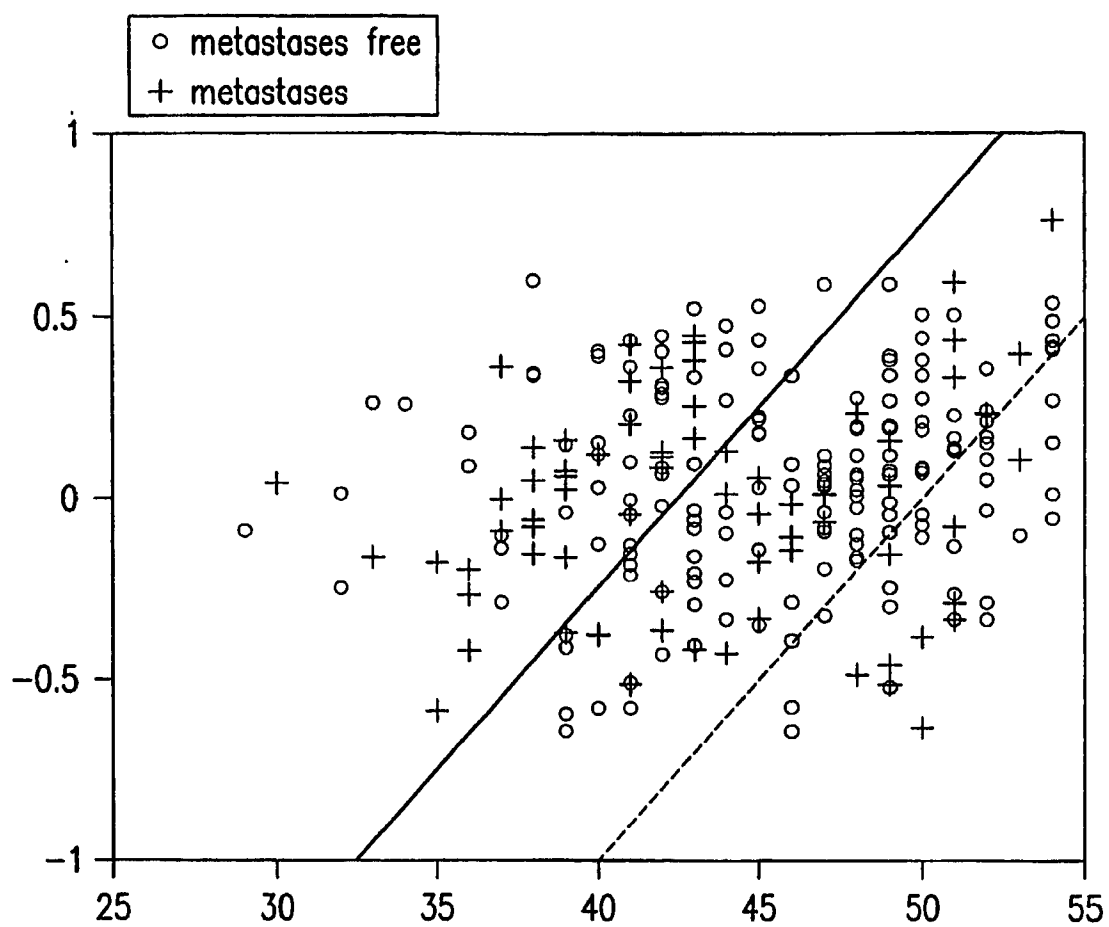


FIG.2A

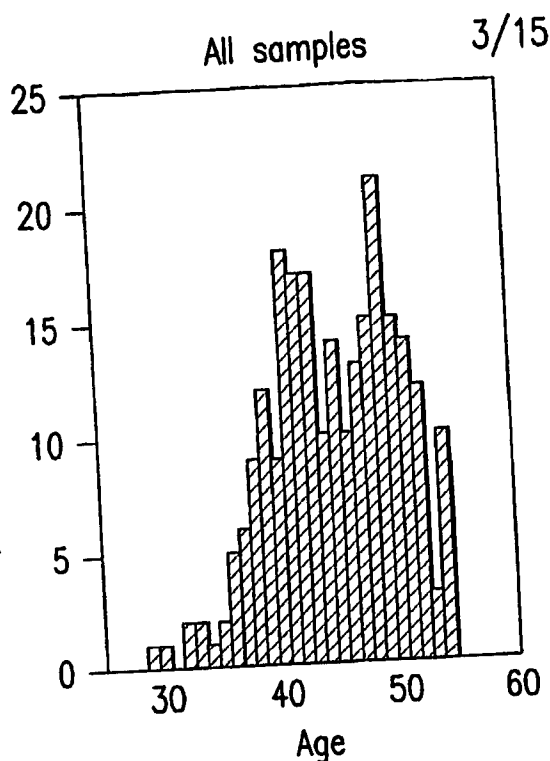


FIG.2B

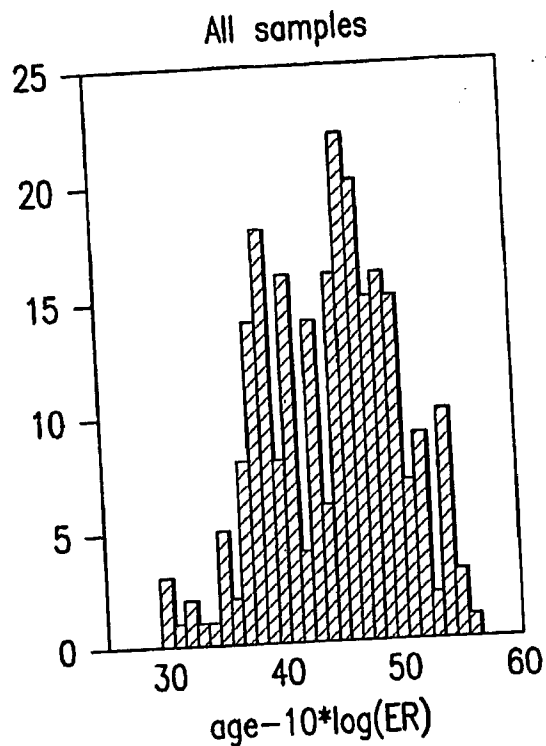


FIG.2C

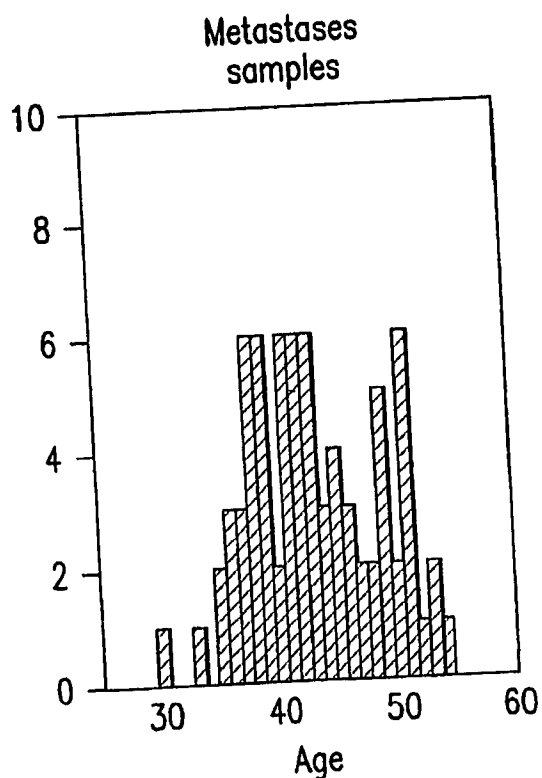


FIG.2D

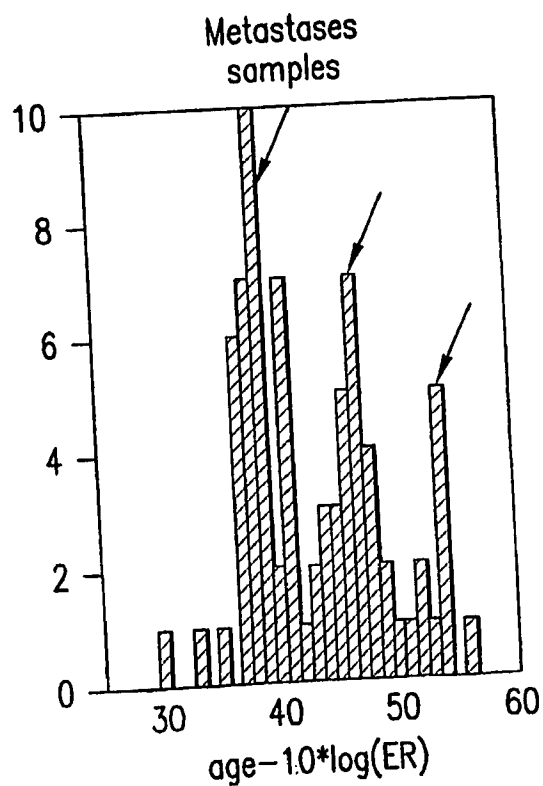


FIG.2E

4/15

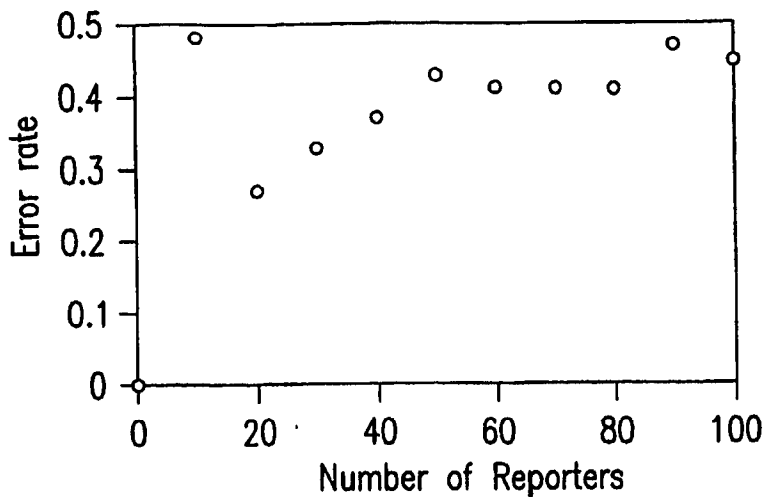


FIG.3A

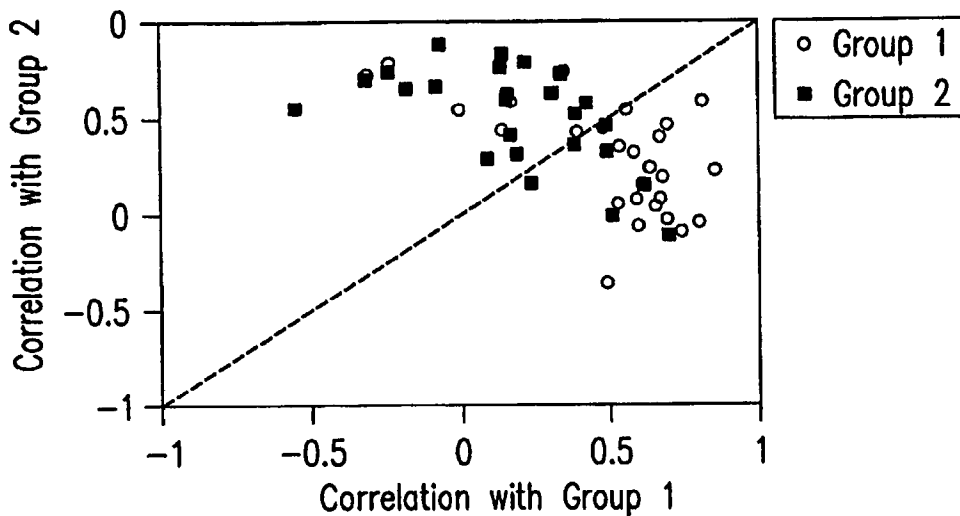
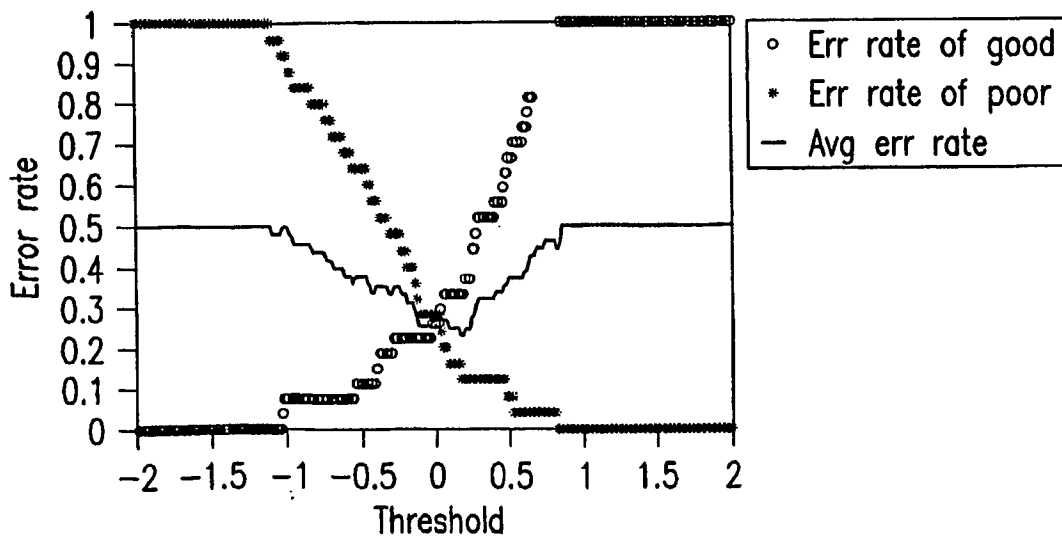


FIG.3B

FIG. 3C
SUBSTITUTE SHEET (RULE 26)

5/15

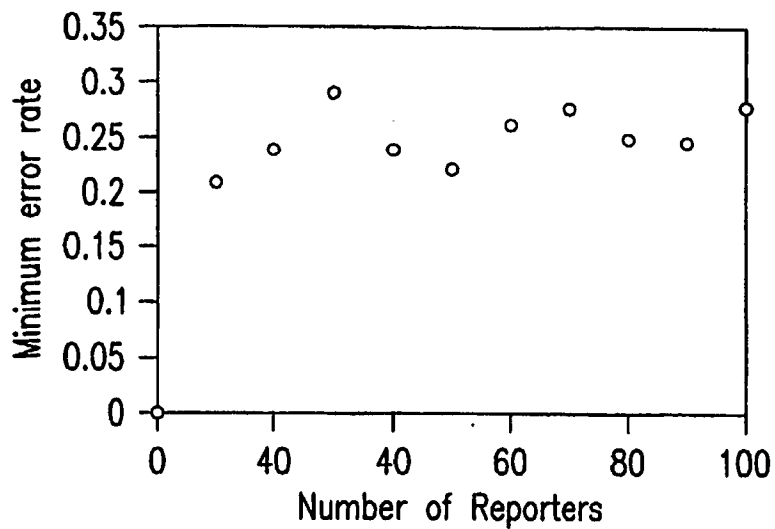


FIG.4A

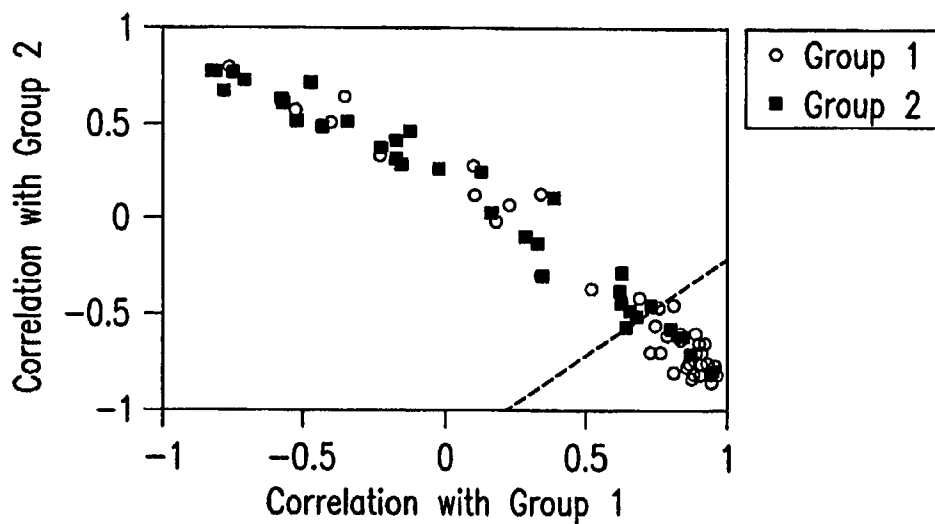


FIG.4B

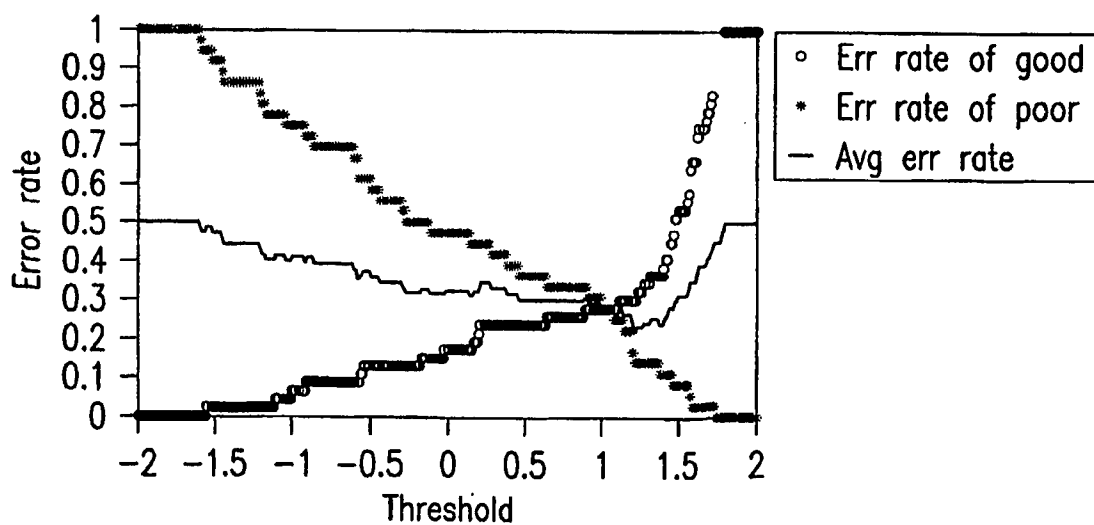


FIG.4C

6/15

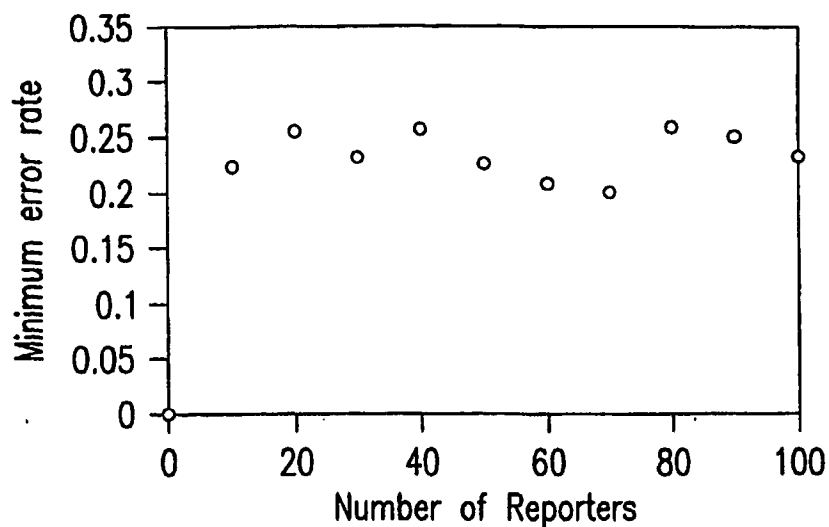


FIG.5A

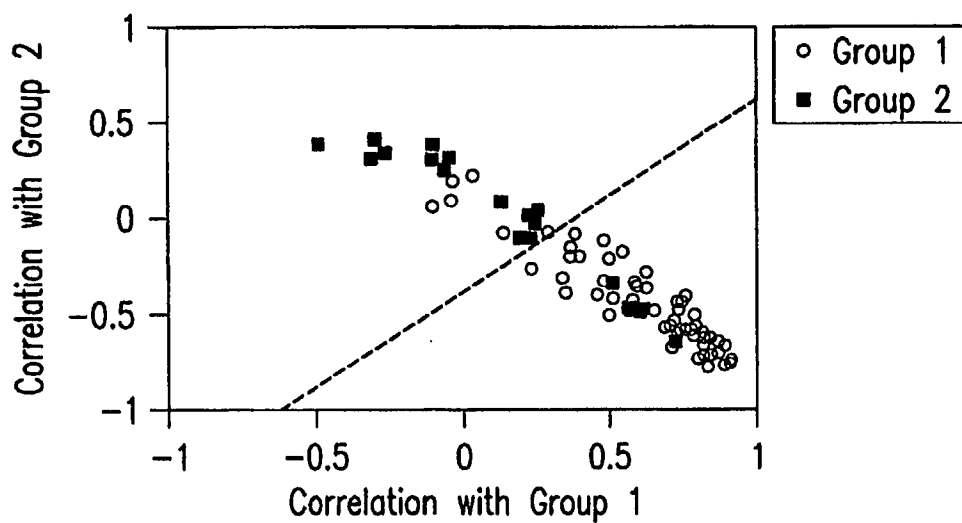
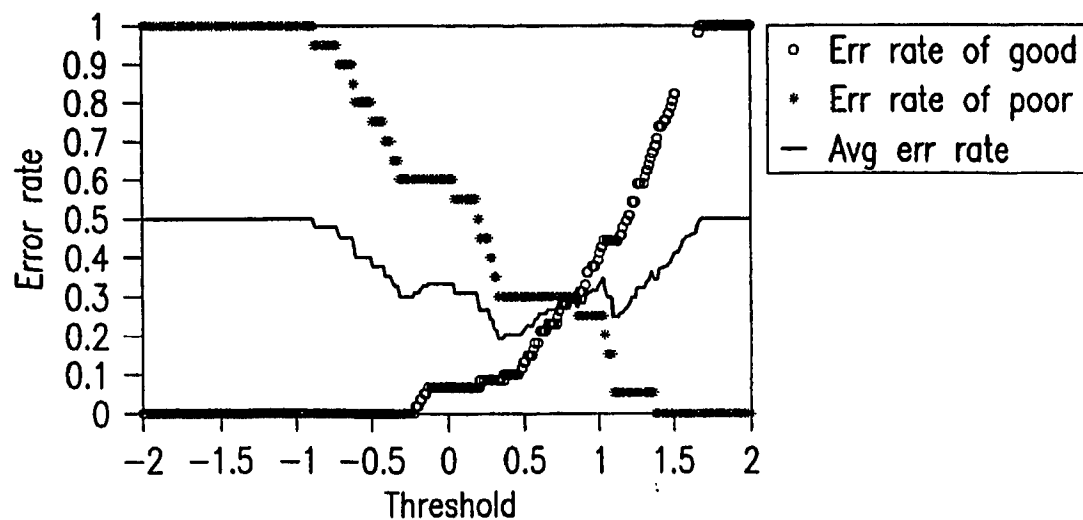


FIG.5B

FIG.5C
SUBSTITUTE SHEET (RULE 26)

7/15

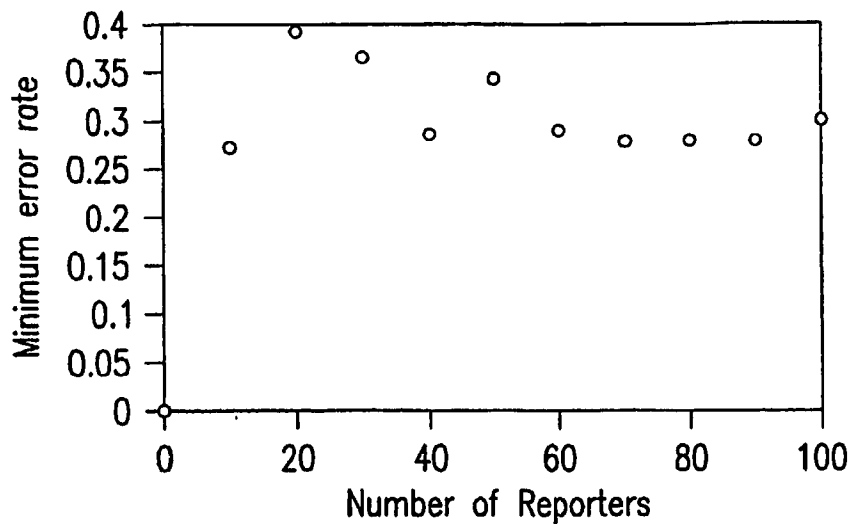


FIG.6A

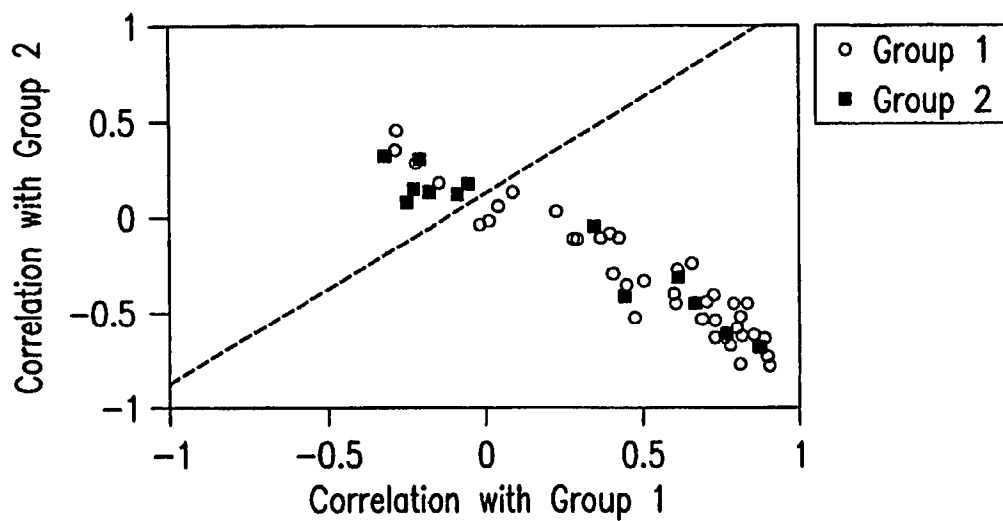
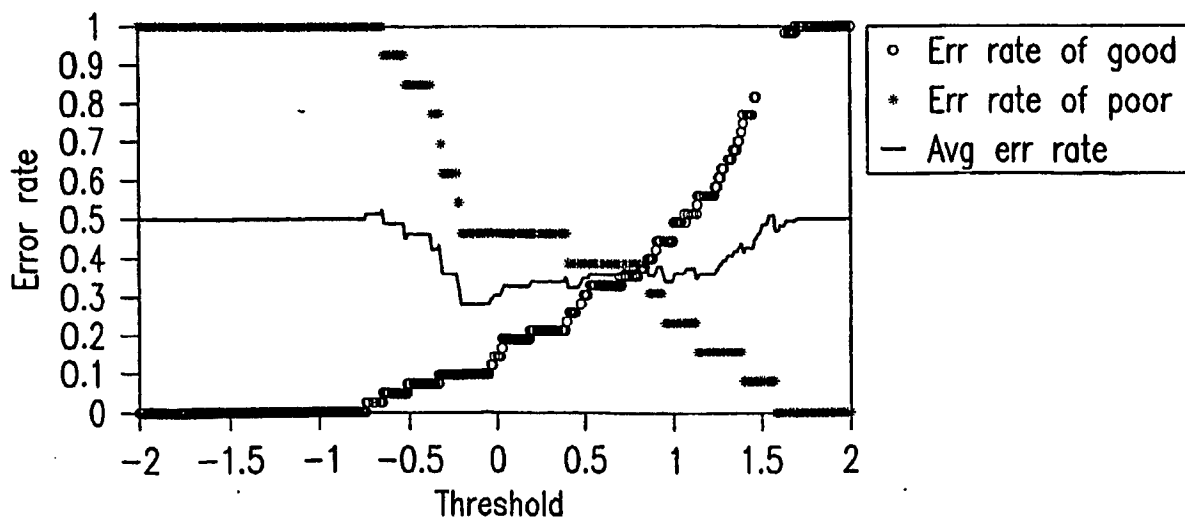


FIG.6B

FIG.6C
SUBSTITUTE SHEET (RULE 26)

8/15

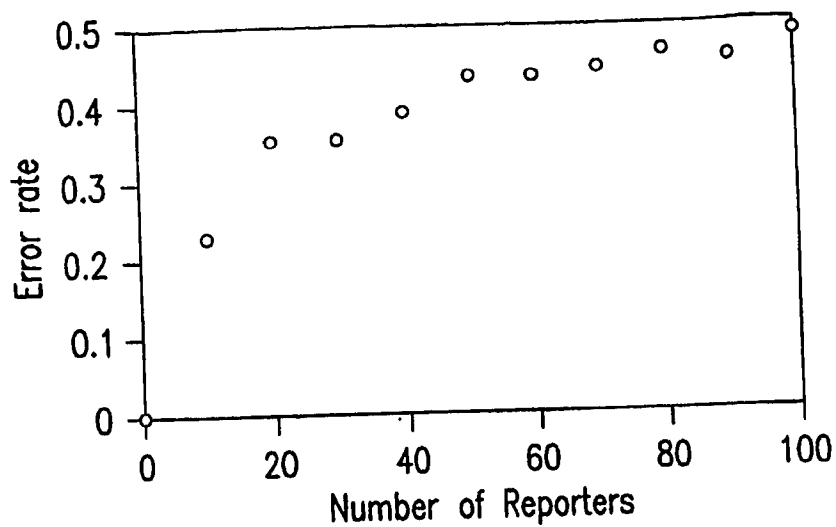


FIG.7A

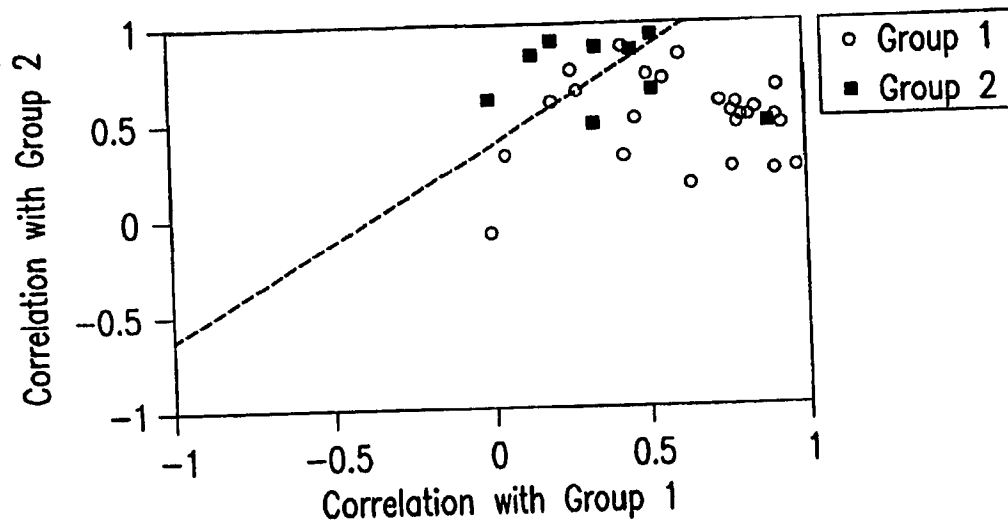
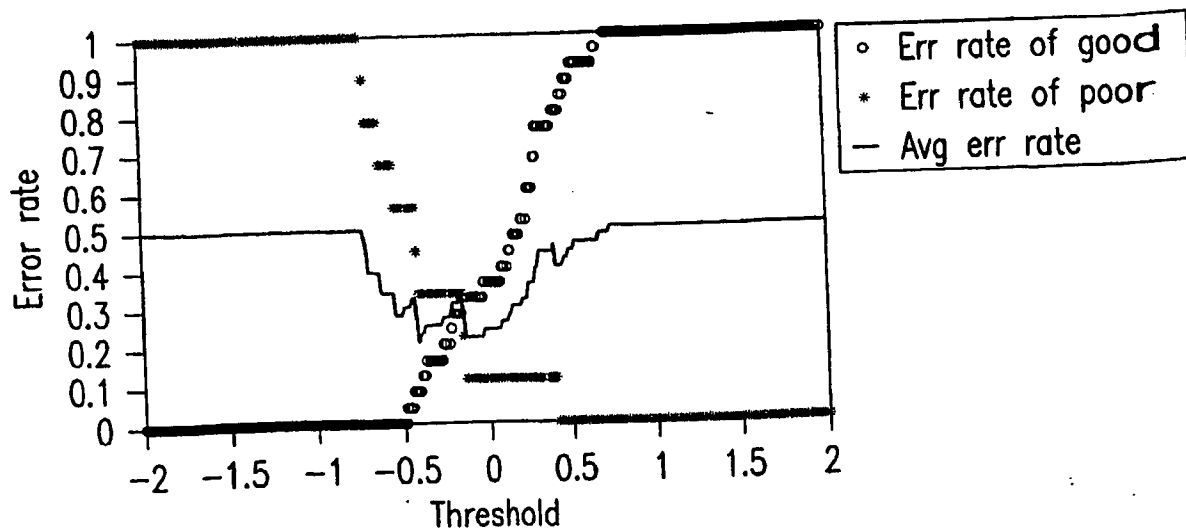


FIG.7B

FIG.7C
SUBSTITUTE SHEET (RULE 26)

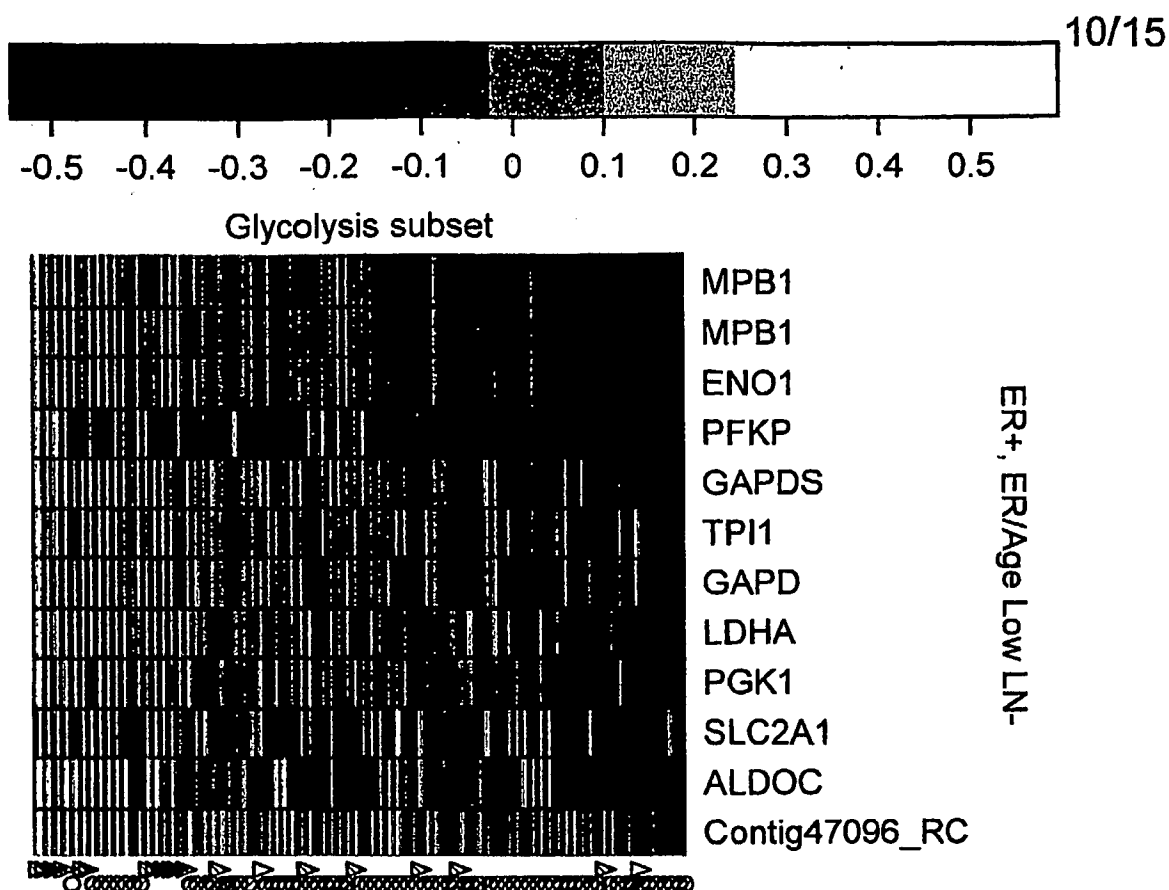


FIG. 8C

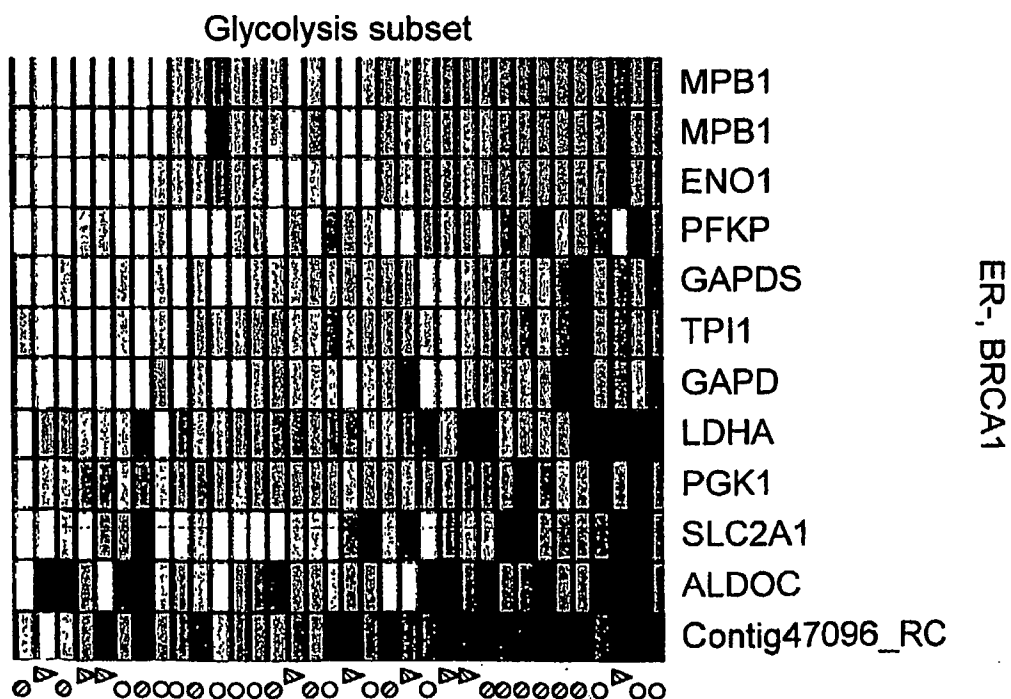


FIG. 8D
SUBSTITUTE SHEET (RULE 26)

11/15

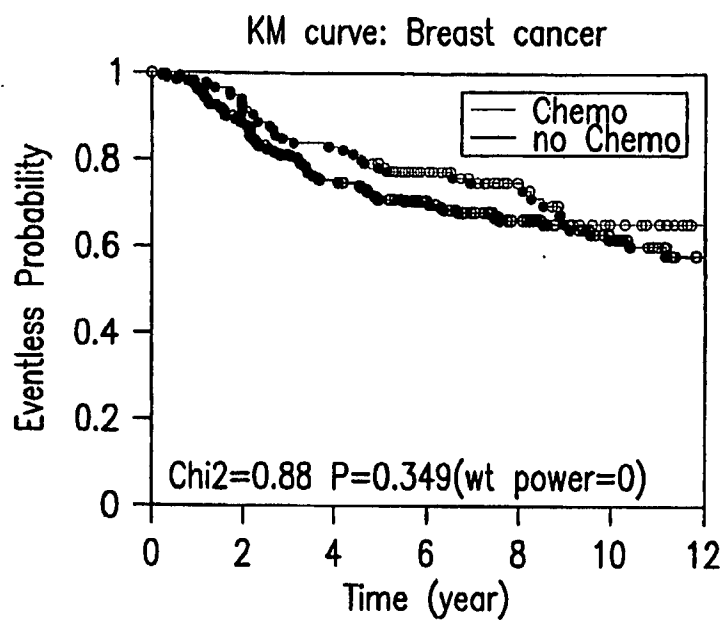


FIG.9

12/15

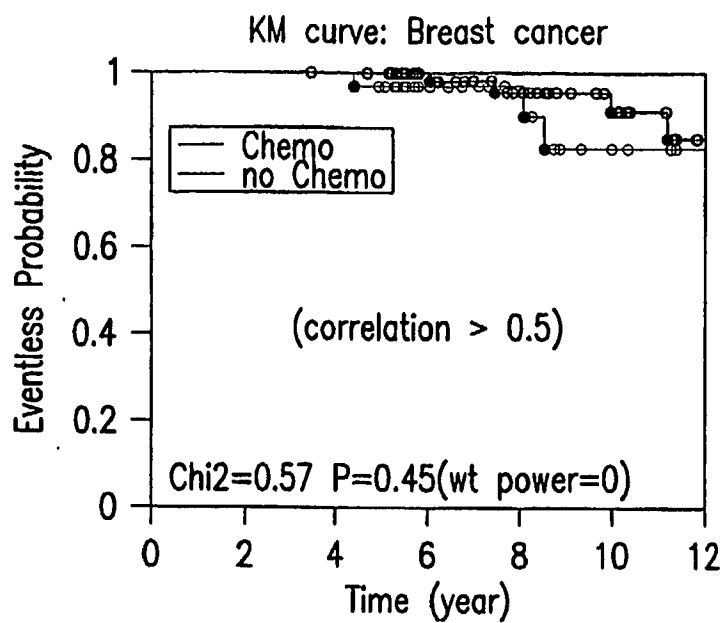


FIG. 10A

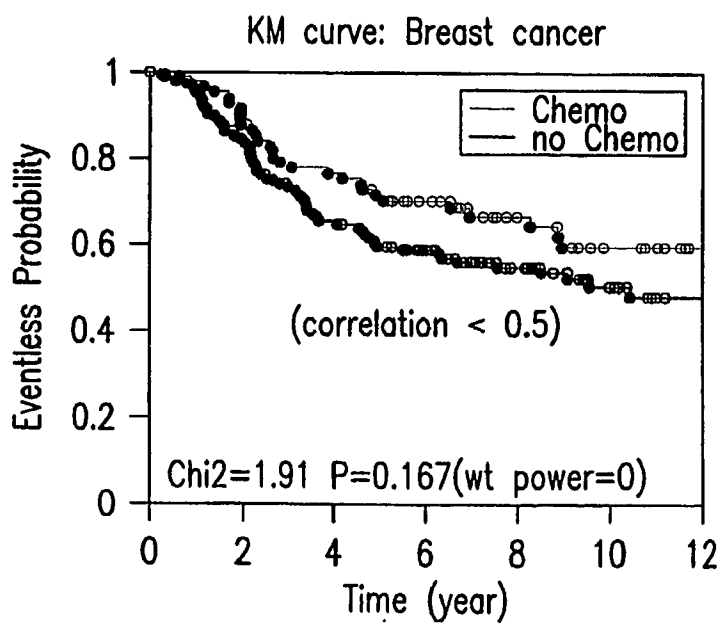


FIG. 10B

13/15

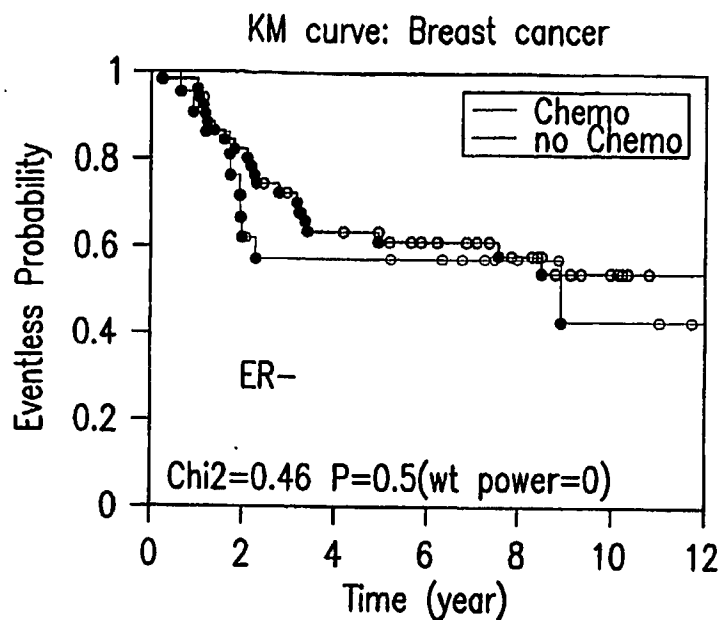


FIG.11A

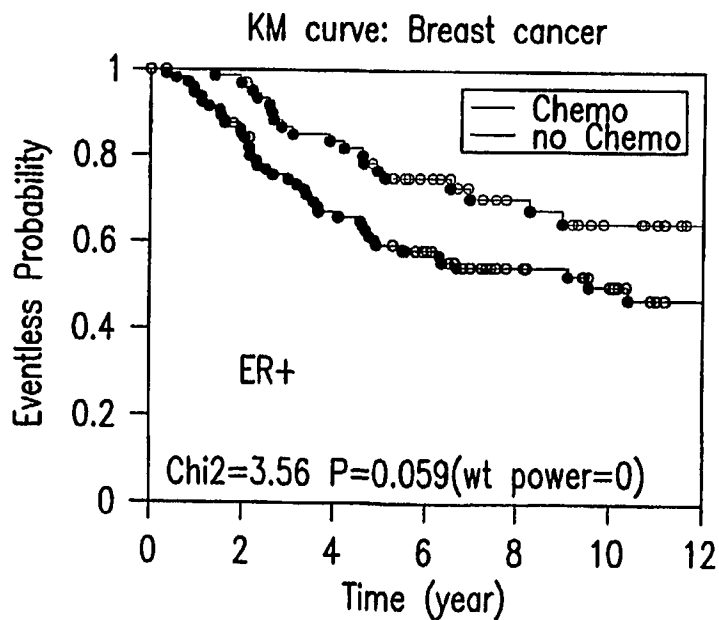


FIG.11B

14/15

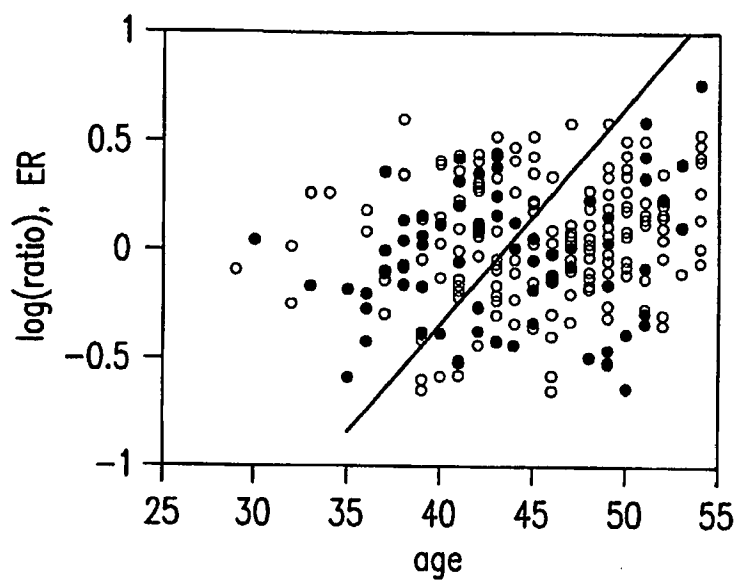


FIG.12

15/15

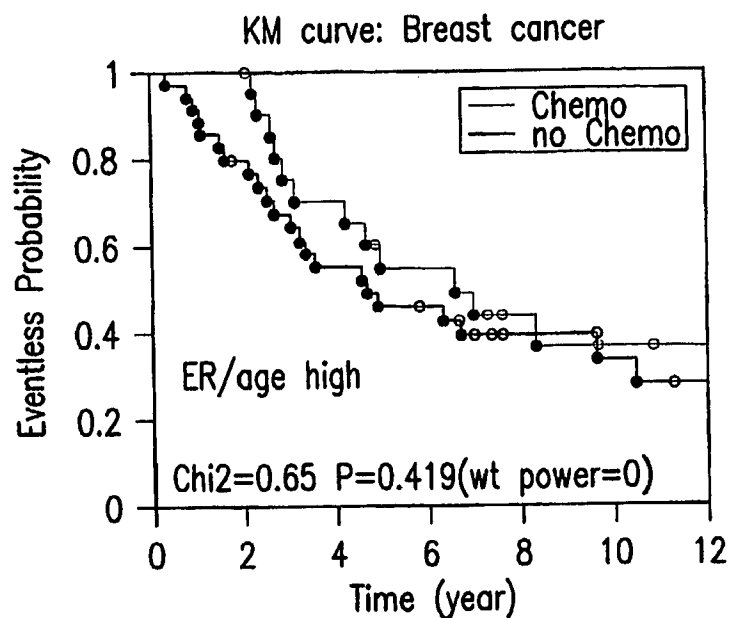


FIG.13A

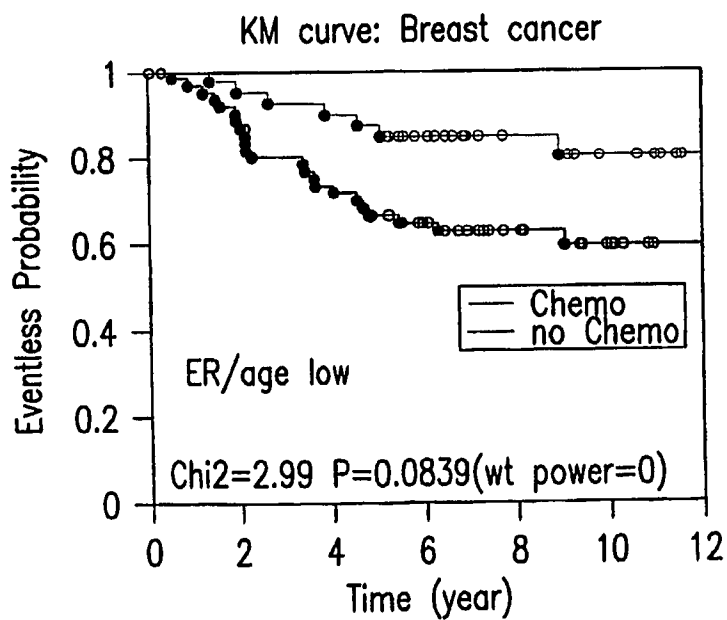


FIG.13B

SEQUENCE LISTING

<110> Rosetta Inpharmatics, LLC

<120> Methods of Predicting Chemotherapy Responsiveness In
Breast Cancer Patients

<130> 9301-249-228

<140>

<141>

<150> 60/650,365

<151> 2005-02-04

<160> 425

<210> 1

<211> 4946

<212> DNA

<213> Homo sapiens

<300>

<308> AB032969

<400> 1

cagcctcagc	ccccagatga	agatggggat	cacagtgaca	agaagatga	acagcctcaa	60
gtggtggttt	taaaaaaggg	agacctgtca	gttgaagaag	tcatgaaaat	taaagcagaa	120
ataaaggctg	ccaaagcaga	tgaagaacca	actccagccg	atggaagaat	catatatcga	180
aaaccagtc	agcatccctc	agatgaaaa	tattcaggtt	taacagcaag	ctcaaaaaag	240
aagaagccaa	atgaagatga	agtaaatacag	gactcgggtca	aaaagaactc	acaaaaacaa	300
attaaaaata	gtagcctcct	ttcttttgac	aacgaagatg	aaaatgagta	agtgtaaata	360
ttttgaattt	agtctacttt	gaaagtatat	ggagtgttca	ttaaaatcac	attttttcct	420
attataaaga	tactacaagt	tctttataga	aagtttagga	aatagagaaa	aaaattta	480
aaactacatc	tattcatcaa	taccctctg	acttaaaatg	ccaactctat	agaaattagc	540
tagtattaac	attttgttat	ttcccttggtg	tggttgtata	tatatgtaaa	ttatatTTTT	600
aagcaaaata	cattttttgt	gtgtaaacaa	aattttataa	atacaactgt	attgcaaagt	660
ttctttgtcc	tgcttctcac	ttgacattgc	attatgagta	ttcttccagg	tcagtaaatt	720
tcaaaaacct	gacattaata	gctacagata	atttcataaa	catctcattg	tatctttttc	780
attagcaata	gctccacttt	gggtggggga	gatgataatg	tgcttgttta	aaaatacctc	840
cccaactcct	gctaagggtg	gccatgagac	tcagctctgg	caagttaaga	aatacaggtg	900
gaattctgct	tgataaagct	gctgggtttt	ttgttacaaa	aggacagact	tggcaa	960
gagcctttgc	tcttatcttt	tcctcctact	tgaggtgcag	agataaaaacc	tgagtaccag	1020
agccactttt	aggcataagg	aaggcagcca	tgtgctttgg	gtcatgttag	taaaaagact	1080
cagagcttgg	ctccttgctg	acatgcctgg	aggagctgct	acaccagctt	ggattgctga	1140
cctctgactt	cttggtagtg	agaagaataa	acactgtgct	taattaggcc	ttggtcaggt	1200
ttcttttata	tgagccaaa	tgagtccta	agtaatacaa	taaataactg	gtcaaactgt	1260
tactggtgga	gggtgtccag	gttcttgcca	ttttggacaa	ataattgaac	aaaacgcaca	1320
aagcaatgaa	tatcctctag	aggtttgcca	ttggttactt	ggcgtacacc	ctgtgtaaat	1380
gaagttagtg	cccgtagcct	gtctgattgg	tgagaaaagt	gaccaatcag	aggctgaagt	1440
gaagttacaa	agtttactc	ctgtgtaaat	gaggacttgg	cctatgacca	gtctgattgg	1500
ttgcaggagg	ggaccaatca	gaggcacttt	catttttcat	ctgcaatgca	gaaaaggcaa	1560
ggggattgca	aaggagtag	cctctgatcc	ttttgttact	taggtatgga	gaggtggggt	1620
tttcttttg	attcagttct	aggaagtcaa	tgtgaatcag	ccttaggttc	cctgtctcca	1680
gaccctatc	tctgcctca	ttttccctc	gagagacgtg	atcctcgtaa	atctttatgg	1740
gaggctgaga	gactgagggt	ctttcttctg	taactgcttc	atgctaactt	gggacacagt	1800
ccctacctat	tgagatcac	gtaactctca	ccctgctttg	tctaggggag	acagggtagc	1860
ttcttgatgg	cgggtggtgt	cttctcctga	aactggctag	aaatcttgtc	acatgatcat	1920
ctaacttgg	ggtctctagg	caaaaggaaa	tggatttgg	taaaagattt	aacagatatg	1980

```

gtccaaaaac caaggcaaat ataatcatta ataatgggct ggccaagggga gggagccatg 2040
aaacccaact tagtgccctt taggtgcccc agctgttgct atattttaga ggcccagtoa 2100
gctagttttc aggtggtgtc ccttactaat cctgattggt tgacatcaaa acagcattct 2160
tcttctagga aaatacataa gccacctgtt tcagcagtta ggagatctag tccccctoga 2220
ttttgcaaag cgaccactgc caaggagcct atccgaattt gtaagggtgac aatactttga 2280
gcaatgttat ccaggctttc cataaaatcc ttggacaagc gttggtaata ggatagggaa 2340
gttgcaatcc cgctaactcc cattcctacc tctgctgtta ttctagccg ttgtgtctgg 2400
tggttgagct taaaggtata atgagggatt ggttggtggg agctatatta atttagggac 2460
atacaatatt tctgtctcca gtctaccact tccaccaaag acaaatcaca gcagaaccga 2520
cctaacttca aaataaactg cagtcccata tactgggcct gattaccocac acaaagtga 2580
acaagaatca ttgtccatat agactctcct agattggcct tgctagaaca tttcacagg 2640
ccatttcagt caaagtcctg agaaagtaac cggtttcaat tgtgccctat tacaaaagaa 2700
aacgtggtta ttaactttat acagacaaat gccatgaatt aagaatattc ataaatagtt 2760
tacaaattct ggagaaatta gaatactcaa tactattaaa gtgtatttca aggctataaa 2820
tagctcaaaa taaaaagatt attcagactc tgaaaaaaca aaaagaagta gcaatatttc 2880
aaacaacaaa agccatacaa attatttcag tcttccatta gttcatttca gtccatgtaa 2940
tcaactcctg ctctacttca tattcatctt tatgaacaca tcagcctttc aattagtgc 3000
ttggaagttt tctgtctaat ccaatggcac actctccaaa gttaccagaa acctgattc 3060
aagagttctt ttcattgaact ccaaagaagt aagccttgga ctgtagctga ttataagtca 3120
cttttttttt ttgagaagga tcaaagcaaa acatcaatta tggatgacaa agtcttaag 3180
acagccataa agacacagtt gacaaatgtg gctatttctg tggcttaca caatttaaca 3240
taatcattac aacatatatt aagacatatt agaattttag aactctcata caatcctgga 3300
acacatatta acaacaatc tctatcagta taacccaaag gaagctaaac accacctcac 3360
acttgacaat gtttccctga taattcaaac attacaaata agcctaatat aagcctaata 3420
tgtcactctt gaacttcagg aagcctaata tccaaaaagt tagtttaagg tcaaaaagtt 3480
ttgaattaac ttttttccat tagtatggtc atatctttct tactaatttg taagttatgt 3540
aatttatcaa tttttttttg ttgttctgtt tcccaacctc tatgtcagat aaagaatcac 3600
ccaggccaaga cacagtggct catgcttgtg gtcccagcac tttgggaagc caagtgagg 3660
gaattgcttg aagccaggaa tctgagccca gctgggcga caaagcaata cccctatctc 3720
tacaaaaaat aaaaaatagc caggtgtggc gacacacacc tgtggtccca gctgctcggg 3780
aggctgagcg ggaggtggc ttgggcccag ggttcaacg ctgcagttag ctgtgattgc 3840
gccactgcac tccagcctgg gcaacagagt aagaactgtc tcaaaaaaaa taaaaaatag 3900
aaataaattt taaaaaaga attaccata ttctctttgt ttttgtttat tcacattaac 3960
ctttattcta tctggaattt atttgagtat acttttttct caaataatca attgtcctag 4020
aaccatgtgt ttctcattta tttgaaaggc catctagtga gagatttctc caaatgttgg 4080
ggtagggaag ggagggaag cactttaaag tctgagcctt tagagggtgat tcctcaagac 4140
cctgcttaat cctaacaatt ttccctcatta gtaaaagtca gcccaaaactg ggggcttgtt 4200
aagatcctta ccagccacat ccactgtgaa ttatgaattt caaagtatct taaaaattg 4260
gtgccacatt atctttttta agtttgtttt gttttgtttt tttgagacag agtctcgtc 4320
tgtcaccggt gctggagtgc agtggcgga tctcagctca ctgcaagctc cgcctcctgg 4380
gttcacacca ttctcttgcc tcggcctccc aagtatctgg gactgcagtc gcccgccacc 4440
acgcccggct aatttttttg tatttttaat agagacgggg tttcactttg ttagccagga 4500
tggtctcaat ctectgacct catgatccac ctgcctcggc ctcccaaagt gctgggatta 4560
caggcaggag ccaccgcgcc tgggcctttt tttaagtttt aagtacctat aaagaacact 4620
gaaaggtgat gtgtgtggat gagctaggaa gacctgaaat aggtctctct taaattaatc 4680
aaattaatcc tgaagccatt ctgcaatact gtctttaatg tatactcact tgttatagaa 4740
gccagggttt tttcccttaa tttgtatcat tgctatatgt gttattgtac caaactacac 4800
tgttttaatt gctgtaaatt ttaatatgtc ttagtatctg ggtgtgggaa tcttgaaagc 4860
atggagtttg tgttattcac cactgtattc tcaaatatca gaagagtatc tggcctacta 4920
agtgcacaat aaacatagtt aaaatg 4946

```

<210> 2

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AB032969

<400> 2

```

taatcctgaa gccattctgc aatactgtct ttaatgtata ctcaattggt atagaagcca 60

```

<210> 3
 <211> 1007
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AF005487

<400> 3
 gaatacagaa tgtgggcaaa ctgcttctg tgccggccgc cagaagggtt gctgagggca 60
 atcactccct ggtgccgggc tccttgaggt tatgcaactg gacatctaga gcctattggt 120
 tgaggaaatgc agtcttgcaa gcctgctctg gatcaagcca cagactgaaa ccccccgaa 180
 gagcaagcac gtttcttgga gcaggctaag tgtgagtgtc atatcttcaa tgggatgaag 240
 cgggtgcagt acctgaacag atacatccat aaacgggagg agaacctgcg cttcgacagc 300
 aacgtggagg agttccaggc agttacggaa ctggggcggc ctgtgcgaga gaactggaac 360
 agccagaagg gcatcccgga ggagaagcgg gacaagatgg acgactactg cagatacaat 420
 tacgggggtt tttgagagct tcacagtgcg gccgcgagtc catcctaagg tgactgtgta 480
 tcctgcaaaag acccagcccc tgcatacccg caacccccctg gtcggctctg tgagtgggtt 540
 ctatccaggc agcattaaag tcagggtggt ccagaatggt caggaagaga aggtgcggt 600
 ggtctccata ggcctgatcc agaattggaga ttggaccttc cagaccctgg tgatgctgga 660
 aacagttcct cggagtggag aggtttacac ctgccaagtg gagcatccaa gcgtgacgag 720
 ccctctcaca gtggaatgga gtacacggac tgaatctgca cagagcaaga tgctgagtgg 780
 agtcgggggc tttgtgctgg gcctgctctt ccttgggaca gggctgttca tctacttcag 840
 gaatcagaaa ggacactctg gacttcagcc aacaggactc ctgcgctgga ctctgagct 900
 gaagtgcaca tgaccacatt caaggaagaa ccttctgccg cagctttgca ggatgaaaag 960
 ctttccact tggtctttat tcttccacaa gagctctctc aggacca 1007

<210> 4
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AF005487

<400> 4
 tttgcaggat gaaaagcttt ccacttggc tcttattctt ccacaagagc tctctcagga 60

<210> 5
 <211> 3200
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AF026941

<400> 5
 caggaagggc catgaagatt aataaagatt tggactcagg gcaaatattt acttagtagc 60
 aataactcaa agaattactg ttgaataaat aagccaatta agcagccaat cacgtactat 120
 gcggatgcac acaaatgaaa ccctcacttc aacctgaaga cattcgacac tgagttacgt 180
 agagggacct gcaggaagcg gtagagaaaa cataaggctt atgcgtttta ttccacacc 240
 aatttcagga tctttgtcac tgacagcagc actaagactt gttaacttta tatagttaag 300
 aagaacaagg ctgagcgcga tgactcacgc ctgtaagcct agaactttgg gaggccaaag 360
 caggcagact gcttgagccc aggagtcca gaccagcctg ggcaacatgg caacacccca 420
 tctctacaaa aaaatacaag aatcagctgg gcgtgggtgat gtgttctgt aatctcagct 480
 actcgggagg cagaggcagg aggattgctt gaacccggga ggcagagggt gtagttagcc 540
 gagatctcgc cactgcactc cagtctggac gacagagtga gactcagctt caaataaata 600
 aataaataca taaatataag gaaaaaata aagctgcttt ctctcttcc tctcttttg 660
 tctcatctgg ctctgctcca ggcactctgc acaatgtggg tgcttacacc tgctgctttt 720
 gctgggaagt tcttgagtgt gttcaggcaa cctctgagct ctctgtggag gagcctgggc 780


```

ccgctgttct gctggctgag ggcaaccttc tggctgctag ctaccaagag gagaaagcag 840
cagctgggtcc tgagagggcc agatgagacc aaagaggagg aagaggaccc tcctctgccc 900
accaccccaa ccagcgtcaa ctatcacttc actcgccagt gcaactacaa atgcggcttc 960
tgtttccaca cagccaaaac atcctttgtg ctgccccttg aggaagcaaa gagaggattg 1020
cttttgctta aggaagctgg tatggagaag atcaactttt caggtggaga gccatttctt 1080
caagaccggg gagaatacct gggcaagttg gtgaggttct gcaaagtaga gttgcggctg 1140
cccagcgtga gcatcgtgag caatggaagc ctgatccggg agaggtgggt ccagaattat 1200
ggtgagttat tggacattct cgctatctcc tgtgacagct ttgacgagga agtcaatgtc 1260
cttattggcc gtggccaag aaagaagaac catgtgaaa accttcaaaa gctgaggagg 1320
tgggtgtagg attatagaat ccctttcaag ataaattctg tcattaatcg tttcaacgtg 1380
gaagaggaca tgacggaaca gatcaaagca ctaaaccttg tccgctggaa agtggtccag 1440
tgctctctaa ttgaaggtga gaattgtgga gaagatgctc taagagaagc agaaagattt 1500
gttattgggt atgaagaatt tgaaagattc ttggagcgcc acaaagaagt gtcctgcttg 1560
gtgcctgaat ctaaccagaa gatgaaagac tcctacctta ttctggatga atatatgctc 1620
tttctgaact gtagaaaggg acggaaggac ccttccaagt ccatcctgga tgttggtgta 1680
gaagaagcta taaaattcag tggatttgat gaaaagatgt ttctgaagcg aggagggaaa 1740
tacatatgga gtaaggctga tctgaagctg gattggtaga gcggaaagtg gaacgagact 1800
tcaacacacc agtgggaaaa ctctagaggt aactggccatt gtctgcaata ctatcccgtt 1860
ggtatttccc agtggctgaa aacctgattt tctgctgcac gtggcatctg attacctgtg 1920
gtcactgaac acacgaataa cttggatagc aaatcctgag acaatggaaa accattaact 1980
ttacttcatt ggcttataac cttgtttgta ttgaaacagc acttctgttt ttgagtttgt 2040
tttagctaaa agaaggaat acacacagga ataatgaccc caaaaatgct tagataaaggc 2100
ccctatacac aggacctgac atttagctca atgatgcgtt tgtaagaaat aagctctagt 2160
gatattctgt ggggcaatat ttaatttgga tttgattttt taaaacaatg tttactgcga 2220
tttctatatt tccattttga aactatttct tgttccaggt ttgttcattt gacagagtca 2280
gtattttttg ccaaatatcc agataaccag ttttcacatc tgagacatta caaagtatct 2340
gcctcaatta tttctgctgg ttataatgct tttttttttt tttgctttta tgccattgca 2400
gtcttgtaact ttttactgtg atgtacagaa atagtcaaca gatgtttcca agaacatatg 2460
atatgataat cctaccaatt ttcaagaagt ctctagaaag agataacaca tggaaagacg 2520
gcgtgggtga gccagccca cggtgcctgt tccatgaatg ctggctacct atgtgtgtgg 2580
tacctgtgtg gtccttttct cttcaaagat ccctgagcaa aacaaagata cgctttccat 2640
ttgatgatgg agttgacatg gaggcagtgc ttgcattgct ttgttcgct atcatctggc 2700
cacatgaggg tgtcaagcaa aagaatagga gtgtagttga gtgctgggt ggccctacat 2760
ttctgagaag tgacgttaca ctgggttgcc ataagatata ctaaaatcac gctggaacct 2820
tgggcaagga agaattgtgag caagagtaga gagagtgcct ggatttcatg tcagtgaagc 2880
catgtcacca tatcatattt ttgaatgaac tctgagtcag ttgaaatagg gtaccatcta 2940
ggtcagttta agaagagtca gctcagagaa agcaagcata agggaaaatg tcacgtaaac 3000
tagatcaggg aacaaaatcc tctccttggt gaaatatccc atgcagtttg ttgatacaac 3060
ttagtatctt attgcctaaa aaaaaatttc ttatcattgt ttcaaaaag caaaatcatg 3120
gaaaattttt gttgtccagg caaataaaag gtcattttta tttaaaaaaa aaaaaaaa 3180
aaaaaaaaa aaaaaggcca 3200

```

<210> 6
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AF026941

<400> 6
 atttttgaat gaactctgag tcagttgaaa tagggtacca tctaggtcag ttttaagaaga 60

<210> 7
 <211> 1799
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AF035284

<400> 7

```

gcttgaaccg gggaggtgga ggttgacgtg agctgagatc acgccattgt actccagcct 60
gggagacaga gcaagactcc atttcaaaaa aaaaaaaaaa aaaaaaaatc cactcatata 120
aaaggtgagc tcagctcact ggtccatttc tcagtggtct ctccatcctc atttgcaaac 180
ctcagaggga taaggcagtt gaacctgatg agcaagaatt ataacagcaa ggaaacatta 240
atgcttagaa ttctgagatc cagcacaaact cagtctgtgg gagctcagct cgctgcccag 300
ggataggtat gacctatgtc tgccttaggc tgctgggaga tgccattctc cagtttcaga 360
agcaggcagg gcaaaggtca agactgtggt attggggctc tttggctctg aaggatcctg 420
gaacctga ttttggttta ttccctccag ggtctaaaga gaacaagagg tgctagctct 480
taccaaaaca gatggtagag agagttgctg gctattttaa aagctctttc atcttttaac 540
tcacctcttc ttttcacctc tttaaccact cctcaggaac agaaccactc taggactggg 600
ggtcttttag ctccataagc aagtgagcag atgggacaag ttagtctttt ctccctagaa 660
acaaagggga tgcccagtg tttcccttg cttcccaacc taaaatttca agtttaataa 720
aatagcaatt agcagaagt accaaattgg gagataatta tcagtcatga ggaaagacac 780
agatttcggt cataaagaat gtaagggtta taagtagaaa ctttctataa cctaaatgat 840
gttatagaat tatttttgag caggagcaga aagattaaat atgatcactt catacttcta 900
aatcagaaat aggaagatta aaaccacaga acagtttgtg atttctattg ctggtagcta 960
ggtatcttac tctgtccact cttgttcaag tatctaactc ttctggaaac caaataggct 1020
ttagaagaga ttatcctata ttccctatcag tataatacta aaatgtaact ttttaatcat 1080
ctggttttta aaagataaac agtttagccc atctctccag agagcaaaca taggaatatg 1140
actcaggagc ctccctaggc ttatcatcag cctcaccacc cgcttcccc tccaaccac 1200
agcctttgct tccagggtgc aggattacta ctttgccctc tcagcagcat ctactctagg 1260
catattgatc attttagaca ctgggagaag agaacctcaa actaggagga aaagacagag 1320
cctccactta gttttgggag gggatggcag acagtcaagg agatgagcgt cctaaggcat 1380
ggtgggtag ggtcagatgc accaccatg gagaggtttg tcaacacaaa gacatggaag 1440
gttagagggt tgtcaacaaa aagacatgga aggttaggtt tgtcaacaca aagacatgga 1500
agattagagg tttgtcaaca caaagataca ggaagaatgg gctgcagaag atttagatgt 1560
tttccatttg ggcacatttt acttagctgg agaactaggt taaaacagc ctgggttagga 1620
aaattagaag caagctggat gcagtggtc atgcctgtaa tccaacact tttgggaggt 1680
ccaggcagga ggatcacttg ggcccaggag gtcaagcctg cagcgagctg agatcacacc 1740
actgcactcc agcctggggt gatagaacaa gaccctgtct caaaaaaaaa aaaaaaaaaa 1799

```

<210> 8

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF035284

<400> 8

```

caaaaagaca tggaaggtta ggtttgtcaa cacaagaca tggaagatta gaggtttgtc 60

```

<210> 9

<211> 1380

<212> DNA

<213> Homo sapiens

<300>

<308> AF052162

<400> 9

```

gtcaaaggat atttatttat aggccttttt ttttttaata tagaatctga ggctgtttgg 60
gctttgactt aaatttccat caggcctctc tccagcaggt aatccctctc cttccgctgg 120
gtcccctggg gaggtgtgaa ctcaagggcc tagcccaaaa acactttttc tgcttttctt 180
aatccttttc cagtcccctc tttttttata aacgttgga gtttgatgtt tctgtttcgg 240
cataacgtaa tccatttcac tgtagcctaa actccagtc gaggttggtt attgttcaaa 300
tgagcagggc ccgagctgga agcgcaaggc agccgcgcgc gtgccgctcc tcccttgccc 360
tcaggccagg tccctgctgg aagcggctgc atcttccctg cagccctggt ttccatgggt 420
actggcgtca cgcagccacc cgagtatggc tgaccttct gcagagagag gagccgcagt 480
cttttgcttg tggaaggaga cgctgggctg tgcggtgcgg aggggtgatga ggatgtctgg 540

```

tgacagccgt	gcggaacacca	ctcctctctg	cagcactgcc	tcccagcgcc	agggctcgcg	600
gcacatccca	ctgagagcgg	gggtcctgcc	ccatcttaga	gtcaaaggca	gaggggcttc	660
caggccctgg	atgggggtatt	ttgggtgcac	ctgaagtccc	tctgacatca	ccttgtttca	720
tcatttttta	tgacagaatt	agaaacccat	ccttcaagca	caataatcat	cacagacttg	780
agtttgcttc	ctaaagcaaa	ggctccgggt	ttgtttggaa	aatttttttg	atttctgaaa	840
tgaattgatt	tttatatttg	gggcatctct	atagaaagtg	accaccaagg	ccagtaagta	900
cgggaaaaaa	tgtttactaa	cttcctcaga	gattcgtgat	acgcgtttct	ccactgacag	960
acatttaaaa	acaaccttca	gctccgtttc	aatcaatcac	ctcgacttgt	tttttagcat	1020
ggacactgcc	agcaggacag	acagggatgg	agtaaaccga	agtcaatttc	agggctcttg	1080
gcgtggttga	cacagaagaa	atcctagtgc	agcctttggt	agctaacagt	cactgatttt	1140
ataattggag	aatgcgtaaa	gattcatttt	tcaaggagaa	gagcctgcaa	atggccaatg	1200
aaggaggtaa	ataaactaag	atattccgag	ggaagggacc	caggccacct	cccttccgca	1260
ggtctgcaga	tgaagggttt	tttgaatgaa	atgccactgt	gcattttcag	aaaaaaaaat	1320
ctctgataaa	cagactttga	atggaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1380

<210> 10

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF052162

<400> 10

cagtaagtac gggaaaaaat gtttactaac ttctcagag attcgtgata cgcgtttctc 60

<210> 11

<211> 1722

<212> DNA

<213> Homo sapiens

<300>

<308> AF055033

<400> 11

ggggaaaaga	gctaggaaaag	agctgcaaag	cagtgtgggc	tttttccctt	tttttgctcc	60
ttttcattac	ccctcctccg	ttttcacctt	tctccggact	tcgcgtagaa	cctgcgaatt	120
tcgaagagga	ggtggcaaaag	tgggagaaaa	gaggtgttag	ggtttggggg	ttttttgttt	180
ttgtttttgt	tttttaattt	cttgatttca	acattttctc	ccacctctct	ggctgcagcc	240
aacgcctctt	acctgttctg	cggcgccgcg	caccgctggc	agctgagggt	tagaaaagcg	300
ggtgtatttt	agatttttaag	caaaaatttt	aaagataaat	ccatttttct	ctcccacccc	360
caacgccatc	tccactgcat	ccgatctcat	tatttcgggtg	gttgcttggtg	ggtgaacaat	420
tttgtggctt	tttttccctt	ataattctga	cccgctcagg	cttgagggtt	tctccggcct	480
ccgctcactg	cgtgcacctg	gcgctgccct	gcttccccca	acctgttgca	aggctttaat	540
tcttgcaact	gggacctgct	cgcaggcacc	ccagccctcc	acctctctct	acatttttgc	600
aagtgtcttg	gggagggcac	ctgctctacc	tgccagaaat	tttaaaacaa	aaacaaaaac	660
aaaaaaatct	ccggggggccc	tcttgggccc	tttatccctg	cactctcgct	ctcctgcccc	720
accccgaggt	aaagggggcg	actaagagaa	gatgggtgtg	ctcaccgcgg	tcctcctgct	780
gctggccgcc	tatgcccggc	cggcccagag	cctgggctcc	ttcgtgcact	gcgagccctg	840
cgacgagaaa	gccctctcca	tgtgcccccc	cagccccctg	ggctgcgagc	tggtaagga	900
gccgggctgc	ggctgtgcga	tgacctgcgc	cctggccgag	gggcagtcgt	gcggcgtcta	960
caccgagcgc	tgcgcccagg	ggctgcgctg	cctcccccg	caggacgagg	agaagccgct	1020
gcacgccctg	ctgcacggcc	gcgggggttg	cctcaacgaa	aagagctacc	gcgagcaagt	1080
caagatcgag	agagactccc	gtgagcacga	ggagcccacc	acctctgaga	tggccgagga	1140
gacctactcc	cccaagatct	tccggcccaa	acacaccgcg	atctccgagc	tgaaggctga	1200
agcagtgaag	aaggaccgca	gaaagaagct	gacccagtc	aagtttgtcg	ggggagccga	1260
gaacactgcc	cacccccgga	tcatctctgc	acctgagatg	agacaggagt	ctgagcaggg	1320
ccctgcgcgc	agacacatgg	aggcttccct	gcaggagctc	aaagccagcc	cacgcatggt	1380
gccccgtgct	gtgtacctgc	ccaattgtga	ccgcaaggga	ttctacaaga	gaaagcagtg	1440
caaaccttcc	cgtggccgca	agcgtggcat	ctgctgggtc	gtggacaagt	acgggatgaa	1500

gctgccaggc atggagtagc ttgacgggga ctttcagtgc cacaccttcg acagcagcaa 1560

cggtgagtga tgcgtccccc cccaaccttt ccttcacccc ctcccacccc cagccccgac 1620

tccagccagc gcctccctcc accccaggac gccactcatt tcattctcatt taagggaata 1680

atatatatct atctatttga ggaaaaaaaa aaaaaaaaaa aa 1722

<210> 12

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AF055033

<400> 12

tccaccccag gacgccactc atttcatctc atttaaggga aaaatatata tctatctatt 60

<210> 13

<211> 1411

<212> DNA

<213> Homo sapiens

<300>

<308> AK001166

<400> 13

aaacaaagag atgccacccc tgtgtgatgg ctttgggtacc cgaacactga tggttcagac 60
 attttcccgt tgcattctgt gttccaagga tgaagtggac ttggatgagt tattagctgc 120
 tagattggta acgtttctga tggacaatta ccaggaaatt ctgaaagtcc ctttggcctt 180
 gcagacctct atagaggagc gtgtggctca tctacgaaga gtccagataa aatacccagg 240
 agctgatatg gatacactt tatctgctcc atcattttgc cgtcaaatta gtccagagga 300
 atttgaatat caaagatcat atggctctca ggaacctctg gcagccttgt tggagggaagt 360
 cataacagat gccaaactct ccaacaaaga gaaaaagaag aaactgaagc agtttcagaa 420
 atcctatcct gaagtctatc aagaacgatt tctacacca gaaagtgcag cacttctggt 480
 tcctgaaaaa cccaaaccga aaccacagct gctaattgtg gcactaaaga agcctttcca 540
 accatttcaa agaactagaa gttttcgaat gtaataatac ttocacagca acaggtgcta 600
 gagaccactg ttgttggttt gagtgaatgg tgggttaggag aaagactttg gtggtggaag 660
 aaagaaaagc ataaaaaaa gactactgaa atatatagaa agattgcctt agtttttaaa 720
 aatgtttggc cattagtatt tttataaaac tcaatgctag ttttaagtgt ataaattggt 780
 taaaatttat gagtcaaata tatagtata atgttaacat gtttgtaatt gctacagaat 840
 ttaagggtat ttttatctct gtgctttctt tttcatgggt tttattaaat aattgtgtat 900
 atacatccta gctactgata tctttattat agccttaaga cttaatttta agtcttaaaa 960
 atagcgtgta tacttgaata agaaagacac tgggtactgt tactgtgatg ctattgactt 1020
 agtagccaat tatcatttct cctgtataaa ttccagtttt tattgtctga cataaatttt 1080
 ttaattgtctt atattgtgat agctatgtct tttattgcag atttatigga tgttatgaca 1140
 gattttacta aagctagtgt tttataaaca tatattatag ttgatgttta cctataagtg 1200
 gagtagattt tcatctgcct gcaatgggtat aatttcagtc ttagctaaaa atggaaagtt 1260
 gaactggata aattctttgg gtacccttag acctctgatt ctaagtcaaa tgcaaatggg 1320
 ttaataaaaa tgagactact tctttataaa atatattttc atccttttga aagtaagtga 1380
 aatgtaataa aacttatttt ttttaaaaat g 1411

<210> 14

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AK001166

<400> 14

acccttagac ctctgattct aagtcaaatg caaatgggtt aaataaaatg agactacttc 60

<210> 15
 <211> 2352
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL049367

<400> 15
 ggcaaacccc ttttaaaatc taatgtctgg gctttgagta ttagctcatt taggggtggac 60
 aaatgcatta ctgttttcaa actgctcaca tttattcagt atttctccaa gttgctatct 120
 actcagcctt atgaatgcc ctcgcttttc taaggccatg tgaaaatcac ggcaactgcc 180
 ttagccttgt gtcactgtct ttttcgttct gcgatatgcc cagttcccaa atcaattata 240
 ggtacctgtt taggagagag gaagatttta cctctcaaag ggtgagattt gaaatttaca 300
 ctaaaaagac aactttacat ttaatgcttc acttaatgag acattccttt ttttataagt 360
 ctatttttct actcagtttc agaacactaa tctgattttc actctgattt ttaacgtttc 420
 tttaaatatt tataatgtag cttctttcaa aatattttca tgaaaaatta cttttattat 480
 accattatgt gcatgttatt ggtagcaggc atagtttatt atttagtact gaaacatgct 540
 cttttaccta acagtaaaca agtatgtttt gatataatc tgtaaatatg cttatagtgg 600
 taagaaatgg acttgaggtc ccaggagatt tcatatttatt caccctgggtc agatacaata 660
 aaggctatga gtataaatac ataacttcct aaccagggtg agggcatgtt catgaatata 720
 aaatcttttg atgctggacc caagagagga aaagttgtag ctaaatgttg atttacttat 780
 aactagacgt ctatgtgaga aaatatatgt atacatatat atgatatgca gaagtcactt 840
 tttttaccag gctttattct ccttacaag ccacagttta actgtctgca acagttgggt 900
 tatgttaatg atagacaaat acccagtggt tgttactttt tccaactacc actgtaatga 960
 taatctttct cacgtatata catgcaactt cttggcttca tttccatgaa gctgtttcaa 1020
 tatattcagt atactttgtc cttaatgctg cttctgttaa cagtgatctc tttctttttt 1080
 tcattcttat atcttcatta gttcatcata aatctgtcca gttgagggct caggaccacg 1140
 gcatgatttc atgactccga agtattttac agaaacattt tttaaataag ggaaatattt 1200
 tatataccag atgtttcaca agtgatggct catagctagt tttttttttt tcttctaaaa 1260
 aatgtcagggt ttttaaaatc atttaccctta ttaaaatgaa aagtgccata cttaactttt 1320
 aaaggaaaaga cctgacttgc tttttctcta ttttagactgt ttttgtaact tactaatctt 1380
 taaactatca ggaaaaaac caaaacttta taccaatgat ttagtaattt tgaggcatag 1440
 ggtagcttac gtagtggagg atgtgccaaa tattctcttc aaatgccacc ttctcaattt 1500
 ataactaaaa tagtgttatc tgactaattc ctctgaattt tgatgtaaga tctatatagg 1560
 cccccaaaat gatcgtagta catgccagtc atttctcagt gaaataaata caataccaga 1620
 gtacattatg gggttttatt ctttctttta tggtagacct gttaatgggg aaaaaataca 1680
 tcaaatcaaa tagaatctta tatctgtatg ttaaaataga gcacttacct gaagtcagt 1740
 gcctggatca tagccctgga tcatttccca gtctgtcctg tgctgtgtga ccttgacaa 1800
 ggcgttcat ctctctgggc ctctatcttcc ccatgtgtaa aacaagtggc tgcagttagat 1860
 gatggctgag agcccttctt gttcccagat gccttgggtc aaagacccca cccctctgct 1920
 ggtcctgcca acgtgttggt gctataagct gcttcagata taaaattgggt ttatctataa 1980
 tgtttgttca ttttaatagct tctaaaaggc ctttttgtaa tacagtgtt tttttctagt 2040
 tttatggact tgattactgt aataatgtct tgttttttagc catgtaacta caaacagata 2100
 ttctcttgat gtcttagtaa atttgcatct gatataatc tgatgagatt ttgttggtat 2160
 gtaatatctt ttggctacgc atctgtccag catcttatta accataatac tgtgatcatt 2220
 atttggaaat atgtcctatg gaaagaataa aagcatgtac ttcacagcta gcatgttcac 2280
 agatttgaag gaagtttcat taaaagcacc attgctttct gtaaaaaaaa aaaaaaaaaa 2340
 aaaaaaaaaa aa 2352

<210> 16
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL049367

<400> 16
 atttggaaat atgtcctatg gaaagaataa aagcatgtac ttcacagcta gcatgttcac 60

<210> 17
 <211> 1130
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL080235

<400> 17
 ggtagcgca cggcgccgt cggcgccgccc gccgccccca ggcgcgcgcg cggcacccgc 60
 gggggcgcca cgcgcgtgcc agcctacccc gccggcgagc cggccggggc gctgtggctg 120
 cagggcgagc cgtgcattt ctgctgccta gacttcagcc tggaggagct gcagggcgag 180
 cggggctggc ggctgaaccg taagcccatt gagtccacgc tgggtggcctg cttcatgacc 240
 ctggtcatcg tgggtgtggag cgtggccgccc ctcatctggc cgggtgcccat catcgccggc 300
 ttcttgccca acggcatgga acagcgccgg accaccgcca gcaccaccgc agccaccccc 360
 gccgcagtgc ccgcagggac caccgcagcc gccgcgcgcg ccgcgcgtgc cgcgcgcgc 420
 gccggcgctca cttcgggggg ggcgaccaag tgaccgcgtc cgtccctccc tgtgtccgtc 480
 ctgtgtccgc gcgcgcgggt gcctttcccg ccggggactc ggccgggtgtg cttcgtgctg 540
 tagttatcgt tagttcctct tcccagatg gggccgcca gagaccccag cgcctttgaa 600
 aagcaagggt tgtgtgcgc ttccagttcc gaaaagcaga tgtttaagcc cttggactga 660
 ggggtgggatc gcagctccga agacggagag gagggaaatg gggccctttc ccctctattg 720
 catccccctg ccgcactect tcccgcacc cagtgccct agattcatgg cagaaaatga 780
 ccaaatcctg tgtatttggg ttatatattt aataactgtt ttaaatgaaa gtttttagtaa 840
 aaaaaataca aaacaaaaag attaaattgc tattgctgta gtaagagaag ctctttgtat 900
 ctgaacatag ttgtatttga aatttgtggg tttttaattt atttaaaatt ggggggaggg 960
 catgggaagg atttaacacc gatatttgt taccgctgaa aatgaacttt atgaaccttt 1020
 tccaagttga tctatccagt gacgtggcct ggtgggcgtt tcttcttgta cttatgtggg 1080
 tttttggctt ttaatacaga cattttcctc caaaaaaaaa aaaaaaaagg 1130

<210> 18
 <211> 60
 <212> DNA

<213> Homo sapiens

<300>
 <308> AL080235

<400> 18
 ctttgaaaag caaggtttgt gctgcgcttc cagttccgaa aagcagatgt ttaagccctt 60

<210> 19
 <211> 2498
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AL137540

<400> 19
 gctgaaacga cagtcttgtc cctgtcagag aaatgacctg aacgaagagc ctcaacattt 60
 tacacactat gcaatctatg atttcattgt caagggcagc tgcttctgca atggccacgc 120
 tgatcaatgc atacctgttc atggcttcag acctgtcaag gcccaggaa cattccacat 180
 ggtccatggg aagtgtatgt gtaagcacia cacagcagc agccactgcc agcactgtgc 240
 ccggttatac aatgaccggc catgggagc agctgatggc aaaacggggg ctcccaacga 300
 gtgcagaacc tgcaagtgtg atgggcatgc tgatacctgt cacttcgacg ttaatgtgtg 360
 ggaggcatca gggaatcgta gtgggtgtgt ctgtgatgac tgtcagcaca acacagaagg 420
 acagtattgc cagaggtgca agccaggctt ctatcgtgac ctgaggagac ccttctcagc 480
 tccagatgct tgcaaacctg gttcctgcca tccagtagga tcagctgtcc ttcttgccaa 540
 ctcaagtgacc ttctgcgacc ccagcaatgg tgactgccct tgcaagcctg ggggtggcagg 600

```

gcgacgttgt gacaggtgca tgggtgggata ctggggcctt ggagactatg gctgtcgacc 660
atgtgactgt gcggggagct gtgacctat caccggagac tgcacagca gccacacaga 720
catagactgg tatcatgaag ttcctgactt ccgtcccgtg cacaataaga gcgaaccagc 780
ctgggagtgg gaggatgcgc aggggttttc tgcacttcta cactcaggta aatgcgaatg 840
taaggaacag acattaggaa atgccaaggc attctgtgga atgaaatatt catatgtgct 900
aaaaataaag attttatcag ctcatgataa aggtactcat gttgaggtca atgtgaagat 960
taaaaaggtc ttaaaatcta ccaaactgaa gattttccga ggaaagcgaa cattatatcc 1020
agaatcatgg acggacagag gatgcacttg tccaatcctc aatcctggtt tggaaatcc 1080
ttagcagga catgaggata taagaacagg caaactaatt gtgaatatga aaagctttgt 1140
ccagcactgg aaaccttctc ttggaagaaa agtcatggat attttaaaaa gagagtgc 1200
gtagcattaa gatggatagc acataatggc acttgtctat gtacaaaaca caaactttag 1260
agcaagaaga cctcagacag gaaactggaa ttttttaaa tgccaaaaca tatagaaatg 1320
tttgaatgca tgggtcttat ctaacttctc tcttctggac ccatgtttta atacagtttt 1380
atttcatgaa gagaaatgaa aacccttaca ctgatctctg ttttctatgg gactgattct 1440
gaaattctta actattaaga atattttaat agcagcatga catttagcag taatccatta 1500
agggcagtac ctctaacaag gacgccttcc agcttcagcg atgttactta cgtttgatgc 1560
tacttaaaag aatgaatgac gttttaagga atccctaacc ctactatcag aaaagggtgt 1620
tgttaaaag ccttctcttg tgtgttacgc atgaactttg gtctgtagg gttaaattgga 1680
acctctccat gtgtatatag tatttccttg tataaagcac tttactacct accacttg 1740
ttgtgaacgt ttggtgactg ctggtgaaag aaggaaaagg gtgtgtgaga aagcctactg 1800
aagcagcagc actgccacta catgtggaca aaagtgacca tataaaagaa gttgtgctat 1860
ttaactctga atacttgag aaactagggt aagatgcaac cagaaaggag aatatgtatg 1920
cgtgaagtct cagctttgag ctggaggcta gattccaaga tgacagccat gatgaaactt 1980
tttaaaaaac taaaccagaa gagacttta aataagagaa agaaatcata aatgtagaca 2040
tatgtctggc taaagggaa atggacttta aattttaaag agctcatttg caatgcactt 2100
gtatacactt caaaaattat ttagacaca gaatttgta tatttttggt cttagtattt 2160
aaacctgaac attgaaacag ttttctcctt tgtctttctt aacagtaata gtcattatat 2220
ttacctgttt tttaacacaa tgtatgtgat agtcaaaaaa tcacagtttt tcattattat 2280
tcactctctg taccacgca taaccactat acatagtttc ttttgactt gaatatacaa 2340
aacatgaaca cagtgcata tgaataattt cacatacaga acctttttt ctctgaagtc 2400
ctgtggactt gcaaatatat atatataatt ctttgttaat ttgtttttat atttcatata 2460
tgtaataaag gaatatgatc tgaaaaaaa aaaaaaa 2498

```

<210> 20

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AL137540

<400> 20

tggaggctag attccaagat gacagccatg atgaaacttt ttaaaaaact aaaccagaag 60

<210> 21

<211> 914

<212> DNA

<213> Homo sapiens

<300>

<308> AL160131

<400> 21

```

cgcaccgcag gagcaacggt tggctctgct gctgtgatgt cgggtgttgag gcccttgac 60
aagctgcccg gcctgaacac ggccaccatc ttgctggtgg gcacggagga tgccttctg 120
cagcagctgg cggactcgat gctcaaagag gactgcgcct ccgagctgaa ggtccacttg 180
gcaaagtccc tccctttgcc ctccagtgtg aatcgccccc gaattgacct gatcgtgttt 240
gtggttaatc ttcacagcaa atacagtctc cagaacacag aggagtccct gcgccatgtg 300
gatgccagct tcttcttggg gaaggtgtgt ttctctgcca caggtgctgg gcgggagagc 360
cactgcagca ttcaccgcca caccgtggtg aagctggccc acacctatca aagccccctg 420

```

```

ctctactgtg acctggaggt ggaaggcttt agggccacca tggcgcagcg cctggtgcgc 480
gtgctgcaga tctgtgctgg ccacgtgccg ggtgtctcag ctctgaacct gctgtccctg 540
ctgagaagct ctgagggccc ctccctggag gacctgtgag ggtggctggc ccctgggctg 600
ccccttctca tggcttcgtg ctgactccat aaacattctc tgttgaggat gtccagtcag 660
ggcttgacag gcccaggctc agcccgcctg ggctgggaag gttccctgca gtgccagtgc 720
tgcagcaggg agagctgggc agaagcagcg agggggccca gctggcgaga ctgtagcccc 780
ctcccactcc cacactcact cttgcagagc ctgtgtcttt aagcagctgg cgtgttacat 840
ctccatttaa ggtttccttt gaacaaaagg tctgtggcta aaaaaagttt aaaaatcact 900
ggtctcattc acca 914

```

```

<210> 22
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> AL160131

```

```

<400> 22
agctggcgtg ttacatctcc atttaagggt tcctttgaac aaaaggctctg tggctaaaaa 60

```

```

<210> 23
<211> 4753
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> D13642

```

```

<400> 23
cttcaatcaa gtagccttcc cactgcagta cacaccagc aaatttgtea tccaccctga 60
gagtaacaac cttattatca ttgaaacgga ccacaatgcc tactactgagg ccacgaaagc 120
tcagagaaag cagcagatgg cagaggaaat ggtggaagca gcaggggagg atgagcggga 180
gctggccgca gagatggcag cagcattcct caatgaaaac ctccctgaat ccatcttttg 240
agctcccaag gctggcaatg ggcagtgggc ctctgtgatc cgagtgatga atcccattca 300
agggaaacaca ctggaccttg tccagctgga acagaatgag gcagctttta gtgtggctgt 360
gtgcaggttt tccaacactg gtgaagactg gtatgtgctg gtgggtgtgg ccaaggacct 420
gataactaaac ccccgatctg tggcaggggg ctctgtggaa gaggtccctg ctgctattgc 480
tggggaaaaa ctggagtttt tgattgggtg ggggaagctg ttgcgtgtct atgacctggg 540
cccattccag gggagggtgt tgattgggtg ggggaagctg ttgcgtgtct atgacctggg 600
aaagaagaag ttactccgaa aatgtgagaa taagcatatt gccaatata tctctgggat 660
ccagactatt ggacataggg taattgtatc tgatgtccaa gaaagtttca tctgggttcg 720
ctacaagcgt aatgaaaacc agcttatcat ctttgctgat gatacctacc cccgatgggt 780
cactacagcc agcctcctgg actatgacac tgtggctggg gcagacaagt ttggcaacat 840
atgtgtggta aggtccccc ctaacaccaa tgatgaagta gatgaggatc ctacaggaaa 900
caaagccctg tgggaccgtg gcttgctcaa gtggggcctc cagaaggcag aggtgatcat 960
gaactaccat gtcggggaga cgggtgctgc cttgcagaag accacgctga tccctggagg 1020
ctcagaatca cttgtctata ccacctgtgc tggaggaatt ggcacctctg tgccattcac 1080
gtcccagtag gaccatgact tcttccagca tgtggaaatg cacctgcggg ctgaacatcc 1140
ccctctctgt gggcgggacc acctcagctt tgcctcctac tacttccctg tgaagaatgt 1200
gattgatgga gacctctgtg agcagttcaa ttccatggaa cccaacaaac aaaagaacgt 1260
ctctgaagaa ctggaccgaa ccccaccgaa agtgtccaa aaactcgagg atatccggac 1320
ccgctacgcc ttctgagccc tcccttcccg gtggggcttg ccagagactg tgtgttttgt 1380
ttccccacc accatcactg ccacctggct tctgccatgt ggcaggaggg tgactggata 1440
attaagactg cattatgaaa gtcaacagct ctttcccctc agctcttctc ctggaatgac 1500
tggcttcccc tcaaatgggc actgagattt gctacacttc tccccacctg gtacatgata 1560
catgacccca ggttccagtg tagaacctga gtccccatt ccccaaagcc atccctgcat 1620
tgatatgtct tgactctcct gtctactttt gcacacaccc ttaattttta attggttttc 1680
ttgtaaatca agttttgtac aatgttatct ctgtgggagg aaggaggcag gctgtggtgg 1740
gaactggtag ggttatgtat cactcctgag ttccactgct ctagaatcta accagaaata 1800
gaaacctagt ttttaagggt actggcatcc atgtgtcttg ttctggagat gaggatgtag 1860

```



```

gtgggaggtt tgaacccaag ttagagcagg aagaactgag tagactcctt ccttccagat 1920
accgacttgg acttgcgcca ctctgtggct cccacacccc aggtctgtgg tggtttcttt 1980
gttttttctt ggttcttttt gctgtgctga tgaacatga cctcaataac catgtgtata 2040
cccacccctc ttccactgg gtattgagga aggggtggctg attcttctct ctcttctact 2100
ctgaggatgt tagtatggg atttttagcat gaattccagc tggggagtct taacagatgc 2160
cccttttact gatagagcac ctaaagcgat ctttggctcc ataggacat aggaaggggtc 2220
agtacagaag aacctagata ctgccctgcc cctgagaact gtgtatatgt ggggcctgtc 2280
tgcagcacc atctcagtg ggttccagag ggcttttagg gtataatgag agcctgttag 2340
gtggaagagg cccagttcca gaaatgttcc agcccccccc tgagaattcc tcctgtttag 2400
ttgtgtggga agccctcgtc ttccaggctg tccttgcgcc ttgaacctgg agaagtgagc 2460
tactgttct caatacttca caaatgtaaa actttctttc gtctgcatgt gctcagccat 2520
ctaaattgag caaatgatct ggtgagcact gggttagaat caggaatggt ggaatacaat 2580
ctgaacctct cagagcccag aacagagggt tcctgacact gtgacactgt ctctggaac 2640
taagtatctc ttgaatcatg acttgggttt agatcagtca agagagaccc aggttttgcc 2700
aggaatcgaa tccctaaata acatgttttt ttctcactta gctcatgaat ttgcatagta 2760
gacagtagtt ctgaattaga ttttgaaaac ctaatttcag ggctcatttt ttctgtggc 2820
cctaaatcca ttctatcaaa ttgtgtgata ctgacatgca gtcatctgag gaactcagcg 2880
tagatacttg agcagctcct cgccctcttt ctaactcaag ttgactaaa atacatacac 2940
tccgtacaga aggtaggggg ttatgtaaga aaggaaaacc taatctatgg aatcaggagt 3000
tgtcaccacc gagcttctc ttggaagtctg cccatcagct tgcttgttct ctgttaagag 3060
gaagggttag gacaaggatt tgggcttgaa tatgtggaaa ggaattttca tagttgttg 3120
tgcaggacct acaaaagttt aaaattagat tggatgtgac tcaatgacaa gtcccactct 3180
tgtaattgtt aaggggacct gattgactcc tgtggtttga ttgagcaacc aggtaaatag 3240
agacctctct ccagctttgg caaaacccat cagaggctgc tgcagaactc agacagaggg 3300
atctgccctt gggtttgctt ccatcctgtt ccattgctaa gcccttgta cttggatct 3360
aggactgaaa agtttttagc tgcctcagct ttcccctgac cttactggca gaggttctgc 3420
agatgtttcc ttggaagat ctcttgccaa gaatagcatt cctttggagg aggggggttc 3480
tagttggaat gttgcttttc ttggttagtg taaatgtatt gctagtgaga cagctgcgg 3540
cgctggaaaa ggctcgtctc acagggagag tgctgggtccc cagaatgtgt gctgttccc 3600
cgctgtgccc tttcttgagc ttgttagagg aaagccagaa aggcattcag atgggatcag 3660
tctggcttcc aaatttttt taattcctaa gttctgtttt attttttaat tttttaaaaa 3720
aaattttatt agagacagtc tctctctctt gctagctgg gagtgcagtg gagtgatcat 3780
agctcactga ggcttgaact cctgggctcg agcaatccac ctccagctcc agagtagggg 3840
agactacaga tgtgtgccac catactcagc tagtttttaa actttcgtag agacagggtc 3900
tccctgtgtt gccaggctg gctcgaact cctgacctca aaaaatcttc ctgccttggc 3960
ctcccagcgc tttgagaggc tgaggcagga ggatcccttg agcccaggag tttgagacca 4020
gcctgggcaa catgacaaaa ccccatctct ccaaaaatac aaaaattggc caggcatggt 4080
ggtgcacact tgtagtccca gtaattaggg ggctgagaca ggaggatcac ttcagcctat 4140
gagtttgagg ctgcagtga ctgtgattgc gccactacac tccagcctgg atgacaggac 4200
gaaacctgtc tcaaaaacac caaaaaacaa aaaccggctc cctgggggtc tggtagcaca 4260
aacgcacatg actgagtgtc caggggttct gaggttgtc cgctgacctg gggctctggc 4320
cctgggagat ctgggggacc tgctgtccta tatgtgatgc tttgaaagaa aggggcatca 4380
ttccaagcca agaggcccca gagagggcac cgtgggggtg tcaggcttct gtgaggcccc 4440

agtgagatcc tgtggctgtg ccccatcacc ctccacccac tctgcccctc cactagctgc 4500
ccaacgatg aatcaacgcc ttggcagagt ttccagcag ggcttgcag agagtgtgtg 4560
tgacctgtgt ggccactgcc ttggggacgg gtgaggagtt agcctggaac attccagcgt 4620
gggcattatt gtctgtttgc aagttcaggg caaaaccagg aatccagttt tgtcgatcca 4680
attgagaaaa catttcatga acaactactt gtggcatgca ttggcactcg gaataaagcg 4740
cactattgtc act 4753

```

<210> 24

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> D13642

<400> 24

aaaccaggaa tccagttttg tcgatccaat tgagaaaaca tttcatgaac aactacttgt 60

<210> 25
 <211> 2591
 <212> DNA
 <213> Homo sapiens

<300>
 <308> D25328

<400> 25
 cccggacgtg cggctccctt cggcctctct gccatggaag cggacgactc ccggggcccc 60
 aagggctcct tgcggaagtt cctggagcac ctctccgggg ccggcaaggc catcggcgtg 120
 ctgaccagcg gcggggatgc tcaaggatat aacgctgccg tccgtgccgt ggtgcgcagt 180
 ggtatctacg tgggggccaa ggtgtacttc atctacgagg gctaccaggg catggtggac 240
 ggaggctcaa acatcgacga ggccgactgg gagagtgtct ccagcatcct gcaagtgggc 300
 gggacgatca ttggcagtgc gcggtgccag gccttcgcga cgcgggaagg ccgcctgaag 360
 gctgcttgca acctgctgca gcgcggcatc accaacctgt gtgtgatcgg cggggacggg 420
 agcctcaccg gggccaacct cttccggaag gagggtgagt ggctgctgga ggagctggcc 480
 aggaacggcc agatcgataa ggaggccgtg cagaagtacg cctacctcaa cgtggtgggc 540
 atggtgggct ccacgcacaa tgatttctgc ggcaccgaca tgaccatcgg cacggactcc 600
 gccctgcaca ggatcatcga ggtcgtcgac gccatcatga ccacggccca gagccaccag 660
 aggaccttcg ttctggaggt gatgggacga cactgtgggt acctggccct ggtgagtggc 720
 ttggcctgcg gtgcggactg ggtgttctct ccagaatctc caccagagga aggctgggag 780
 gagcagatgt gtgtcaaac ctgcggaac cgtgcccgga aaaaaaggct gaattattat 840
 attgtggctg aaggagcaat tgatacccaa aataaaacca tcacctctga gaaaatcaaa 900
 gagottgtcg tcacgcagct gggctatgac acacgtgtga ccacctcgg gcacgtgcag 960
 agaggaggga ccccttcggc attcgacagg atcttgccca gccgcatggg agtggaggca 1020
 gtcacgcctt tgetagaggc caccocggac accccagctt gcgtcgtgtc actgaacggg 1080
 aaccacgccg tgcgcctgcc gctgatggag tgcgtgcaga tgactcagga tgtgcagaag 1140
 gcgatggacg agaggagatt tcaagatgcg gttcgactcc gagggaggag ctttcggggc 1200
 aacctgaaca cctacaagcg acttgccatc aagctgccgg atgatcagat cccaaagacc 1260
 aattgcaacg tagctgtcat caacgtgggg gcaccgcgg ctgggatgaa cgcggccgta 1320
 cgctcagctg tgcgcgtggg cattgcgcac ggccacagga tgctcgccat ctatgatggc 1380
 tttgacggct tcgccaaggg ccagatcaaa gaaatcggct ggacagatgt cgggggctgg 1440
 accggccaag gaggtcccat tcttgggaca aaacgcgttc tcccggggaa gtacttggaa 1500
 gagatcgcca cacagatgcg cacgcacagc atcaacgcgc tgctgatcat cgggtggattc 1560
 gaggcctacc tgggactcct ggagctgtca gccgccggg agaagcaaga ggagttctgt 1620
 gtccccatg tcatggttcc cgctactgtg tccaacaatg tgccgggttc cgatttcagc 1680
 atcggggcag acaccgccct gaacactatc accgacacct gcgaccgat caagcagtc 1740
 gccagcgga ccaagcggcg cgtgttcac atcgagacca tgggcggcta ctgtggctac 1800
 ctggccaaca tgggggggct cgcggccgga gctgatgccg catacatttt cgaagagccc 1860
 ttgcacatca gggatctgca gtccaacgtg gagcaactga cggagaaaat gaagaccacc 1920
 atccagagag gccttgtgct cagaaatgag agctgcagtg aaaactacac caccgacttc 1980
 atttaccagc tgtattcaga agagggcaaa ggcgtgtttg actgcaggaa gaacgtgctg 2040
 ggtcacatgc agcaggggtg ggcacctct ccatttgata gaaactttgg aaccaaaatc 2100
 tctgccagag ctatggagtg gatcactgca aaactcaagg aggccgggg cagaggaaaa 2160
 aaatttacca ccgatgattc catttgtgtg ctgggaataa gcaaaagaaa cgttattttt 2220
 caacctgtgg cagagctgaa gaagcaaacg gattttgagc acaggattcc caaagaacag 2280
 tgggtggctc agctacggcc cctcatgaaa atcctggcca agtacaaggc cagctatgac 2340
 gtgtcggact caggccagct ggaacatgtg cagccctgga gtgtctgacc cagtcccgcc 2400
 tgcatgtgcc tgcagccacc gtggactgtc tgtttttgta acacttaagt tattttatca 2460
 gcactttatg cagttattat tgacattaat acctaactcg cgagtgccca tctgccccac 2520
 cagctccagt gcgtgctgtc tgtggagtgt gtctcatgct ttcagatgtg catatgagca 2580
 gaattaatta a 2591

<210> 26
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> D25328

<400> 26

tattttatca gcactttatg cacgtattat tgacattaat acctaatacg cgagtgccca 60

<210> 27

<211> 2573

<212> DNA

<213> Homo sapiens

<300>

<308> D50402

<400> 27

```

gaatcgccg atgtgaaccg aatggtgatg taagaggcag ggactcggc tgcggatggg 60
taacagggcg tgggctggca cacttacttg caccagtgcc cagagagggg gtgcaggctg 120
aggagctgcc cagagcaccg ctccactcc cagagtacct gaagtcggca tttcaatgac 180
aggtgacaag ggtcccaaaa ggctaagcgg gtccagctat ggttccatct ccagcccgac 240
cagcccgacc agccagggc caggcaagc acctccaga gagacctacc tgagtgaaga 300
gatccccatc ccagacacaa aaccgggcac cttcagcctg cggaagctat gggccttcac 360
ggggcctggc ttcctcatga gcattgcttt cctggacca ggaaacatcg agtcagatct 420
tcaggctggc gccgtggcgg gattcaaact tctctgggtg ctgctctggg ccaccgtgtt 480
gggcttgctc tgccagcgac tggctgcacg tctggcgctg gtgacaggca aggacttggg 540
cgaggctctg catctctact acctaaagg gcccgcacc gtccctctggc tgacctcga 600
gctagccatt gtgggtccg acatgcagga agtcctggc acggccattg cattcaatct 660
gctctcagct ggacgaatcc cactctgggg tggcgtctc atcaccatcg tggacacctt 720
cttctctctc ttcctcgata actacgggtc ggggaagctg gaagcttttt ttggactcct 780
tataaccatt atggccttga ctttggcta tgagtatgtg gtggcgctc ctgagcaggg 840
agcgtctctt cggggcctgt tccctgccctc gtgcccgggc tgcggccacc ccgagctgct 900
gcaggcggtg ggcattgttg gcgccatcat catgccccac aacatctacc tgcactcggc 960
cctggccaag tctcgagaga tagaccgggc ccgcccagcg gacatcagag aagccaacat 1020
gtacttccctg attgagccca ccctcgccct gtccgtctcc tttatcatca acctctttgt 1080
catggctgtc tttgggcagg ccttctacca gaaaaccaac caggctgctg tcaacatctg 1140
tgccaacagc agcctccacg actacgcca gatcttcccc atgaacaacg ccaccgtggc 1200
cgtggacatt taccaggggg gcgtgatcct gggctgcctg ttcggccccg cggccctcta 1260
catctggggc ataggtctoc tggcggtgg gcagagctcc accatgacgg gcacctacgc 1320
gggacagttc gtgatggagg gcttctctgag gctgcggtgg tcacgcttgc cccgtgtcct 1380
cctcaccggc tccctgcgca tccctgccac cgtgctctg gtgtcttcc gggacctgag 1440
ggacttgctg ggcctcaatg atctgctcaa cgtctgcag agcctgctg tcccgctcgc 1500
cgtgctggcc atcctcacgt tcaccagcat gccaccctc atgcaggagt ttgccaatgg 1560
cctgctgaac aaggctcgtc cctcttccat catggtgcta gtctgcgcca tcaacctcta 1620
cttctgggtc agctatctgc ccagcctgcc ccaccctgcc tacttcggcc ttgcagcctt 1680
gctggccgca gcctacctgg gcctcagcac ctacctggtc tggacctgtt gccttgccca 1740
cggagccacc tttctggccc acagctccca ccaccacttc ctgtatggg tccttgaaga 1800
ggaccagaaa ggggagacct ctggctaggc ccacaccagg gcctggctgg gagtggcatg 1860
tatgacgtga ctggcctgct ggatgtggag gggcgcgctg caggcagcag gatggagtgg 1920
gacagttcct gagaccagcc aacctggggg ctttagggac ctgctgtttc ctagcgagc 1980
catgtgatta ccctctgggt ctcatgtgta tcatctgtaa aatggagacg ccaccacctt 2040
tgccatggag gttaagcact ttaacacagt gtctggcact tgggacaaaa acaaacaaac 2100
aaacaaaaaa catttcaaaa ggtatttatt gagcacctgc aggcgtgacc tgacagccca 2160
agggtggttg gggtaggggc ttgaggactt gggcgggaca caggctccaa actggagctt 2220
gaaatagtg ctgatgaatg ttaaattatc tatctatcta tttatttatt tatttgagac 2280
agggaaagg tctccctctg ttgccaaggc tggagtgcag tggcgcaatc ttaactcatt 2340
gcaacctcca ccttctgggt tcaagcatt tctttattc agccccggga gtggcgcgcg 2400
ccaccacgcc cagctaattt gtgtatttcc agcagagacg gggtttgcca tgctggccag 2460
gctggtctcg aactgctgga ttcaagtgat ccgcccctct ccgtctccca aagtgtctgg 2520
aattacaggc gtgagccacc aaaaccgggc ctgattaaag ttaaataaat acg 2573

```

<210> 28

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> D50402

<400> 28

tggagggttaa gcactttaac acagtgtctg gcacttggga caaaaacaaa caaacaacaa 60

<210> 29

<211> 3672

<212> DNA

<213> Homo sapiens

<300>

<308> L27560

<400> 29

```

acatgtgcat atttcattcc ccaggcagac attttttaga aatcaatata tgccccaata 60
ttggaaagac ttgttcttcc acggtgacta cagtacatgc tgaagcgtgc cgtttcagcc 120
ctcatttaat tcaatttgta agtagcgac gagcctctgt gggggaggat aggctgaaaa 180
aaaaaagtgg gctcgtattt atctacagga ctccatatag tcatatatag gcatataaat 240
ctatgctttt tctttgtttt tttctttctt cctttctttc aaagggtttgc attaaactttt 300
caaagtagtt cctatagggg cattgaggag cttcctcatt ctgggaaaac tgagaaaacc 360
catattctcc taatacaacc cgtaatagca tttttgcctg cctcgaggca gaggttcccg 420
tgagcaataa actcagcttt tttgtggggc acagtactgg atttgacagt gattccccac 480
gtgtgttcat ctgcaccacac cgagccaggc agaggccagc cctccgtggt gcacacagca 540
cgcgcctcag tccatcccat tttagtcttt aaaccctcag gaagtcacag tctccggaca 600
ccacaccaca ttgagcccaa cagggtccacg atggatccac ctagtccac cccagccttt 660
ttctttcatc tgaacagaat gtgcattttt ggaagcctcc ctactctcc atgctggcag 720
agcaggaggg agactgaagt aagagatggc agaggggagat ggtggcaaaa aggttttagat 780
gcaggagaac agtaagatgg atggttccgg ccagagtcga tgtggggagg aacagagggc 840
tgaagggaga gggggctgac tgttccattc tagctttggc acaaagcagc agaaaggggg 900
aaaagccaat agaaatttcc ttagcttccc caccatatgt attttcatgg atttgagagg 960
aaagagagga aaatggggga atgggttgca aaatagaaat gagcttaatc caggccgcag 1020
agccaggga ggtgagtaac cttaggaggg tgctagactt tagaagccag ataggaagaa 1080
tcagtctaaa ctggccatgc tttggaaggg acaagactat gtgctccgct gccacacctc 1140
agcctgcaat gagggactga ggcccacgag tctttccagc tcttctccca ttctggccag 1200
tcctctgcac ctccctgggg tgaggatgg aaggaaagct gggacaagca gggaacgcac 1260
gattcaggga tgctgtcact cggcagccag attccgaaac tccattctc caatgacttc 1320
ctcaaccaat ggggtggcct gtgactgttc tttaaggctg aagatatcca ggaaaggggg 1380
cttggaact ggccaaggag acccctcgt gctgtggaca cagctctctt cactctttgc 1440
tcattggcatg acacagcgga gaccgcctcc aacaacgaat ttggggctac gaagaggaat 1500
agcgaaaaag caaatctgtt tcaactgatg ggaaccctat agctatagaa cttgggggct 1560
atctcctatg ccctggaca ggacagtgg ctggggacag gagaagtgt caatcttcat 1620
gagacaaaag ggccgatca aggcagccac aaggccttga cctgccagt cagcatgcc 1680
catctctctc gacagctgtc ccctaaaacc aactcacgtt tctgtatgtc ttaggccagt 1740
atcccaaac tcttccacgt cactgttctt tccaccatt ctccctttgc atcttgagca 1800
gttatccaac taggatctgc caagtggata ctgggggtgc actccctga gaaaagactg 1860
agccaggaaac tacaagctcc cccacattc ctcccagcct ggacctaat cttgagaggg 1920
gctctctctt caggactgt gtctggactt tgagcaggct tctgcccctt gcgttggtc 1980
tttgctgcca gccatcaggt gggggattag agcctggtgt aagtgcgcca gactcttccg 2040
gtttccaaa ttctgtcctg cgaacccaaa cctgtgagtc tcttctgcat gcaggagttt 2100
ctcctgggga gctgtcact cccagagaa gctgggcctt catggacaca tggaactaag 2160
cctcccaaat gggagtctgt gctgagccca ggggtggggag atcctgggaa gggaggcact 2220
ggaggaagac ggcacctctt ccccatggc aggggtgtgag ggaggcaggt ttggaatggt 2280
gcgagtatgg caatctaagc aggggtctgg tctctttgac tccaggctcg ctttggccga 2340
ctgtctgtc acccagagac cttggactcc ggactatcca tggctccgaa tctaagtgt 2400
gccactccc atgtcacac ccacagaagg tcttcccatc ccttttagat tcgtgcctca 2460
ctccaccagt gaggaagatg cctctgtctt tcccacgact gccaggagat aggggaagccc 2520
agccaggact gaccctcctt cctccagcct gccctgacct acctggcaaa gcagggcaca 2580

```

```

tggggaggaa gagactggaa cctttctttg acagccaggc ctagacagac aggcctgggg 2640
acactggccc atgaggggag gaaggcaggc gcacgaggtc cagggaggcc cttttctgat 2700
catgcccctt ctctcccacc ccatctcccc accaccacct ctgtggcctc catggtacct 2760
ccacagggct ggccctcccct agaggggtggg cctcaaccac ctctgtccgc cagcaccggg 2820
ttagtgagac agggctgcca cgcaaccgcc aagccccctt caaggtggga cagtaccccg 2880
gacccatcca ctcactcctg agaggctccg gcccagaatg ggaacctcag agaagagctc 2940
taaggagaag aaaccccata gcgtcagaga ggatatgtct ggcttccaag agaaaggagg 3000
ctccgttttg caaagtggag gagggacgag ggacaggggt ttcaccagcc agcaacctgg 3060
gccttgtact gtctgtgttt ttaaaaccac taaagtgcaa gaattacatt gcactgtttc 3120
tccacttttt attttctctt aggcctttgt ttctatttca aacatacttt cttggttttc 3180
taatggagta tatagtttag tcatttcaca gactctggcc tcctctcctg aaatcccttt 3240
ggatggggaa agggaagggt gggagggtcc gaggggaagg ggaccccagc ttccctgtgc 3300
ccgctcacc cactccacca gtccccggtc gccagccgga gtctcctctc taccgccact 3360
gtcacaccgt agcccacatg gatagcacag ttgtcagaca agattccttc agattccgag 3420
ttgtaccgg ttgttttctg ttgtgtgtgt gttgtttttc tttttctttt tttttttgaa 3480
gacagcaata accacagtac atattactgt agttctctat agttttacat acattcatac 3540
cataactctg ttctctcctc tttttgttt tcaactttaa aaacaaaaat aaacgatgat 3600
aatctttact ggtgaaaagg atggaaaaat aaatcaacaa atgcaaccag tttgtgagaa 3660
aaaaaaaaaa aa 3672

```

<210> 30
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> L27560

<400> 30
 agcaacctgg gccttgtact gtctgtgttt ttaaaaccac taaagtgcaa gaattacatt 60

<210> 31
 <211> 1416
 <212> DNA
 <213> Homo sapiens

<220>
 <221> Modified_base
 <222> 1 ... 1416
 <223> n = a,c,g, or t

<300>
 <308> M55914

<400> 31
 aggaattccg gaattccgga attccgatgg atggaacaga aaataaatct aagtttggtg 60
 cgaacgccat tctgggggtg tcccttgccg tctgcaaagc tgggtgccgtt gagaaggggg 120
 tcccctgtac cgccacatcg cgtacttggc tggcaacttc gaagtcatcc tgccagtccc 180
 ggcgttcaag tgtcatcatc aatggcggtt ctcatgctgg caacaagctg gccatgcaga 240
 gtctgtcctc ccagtcggtg cagcaaaactc aggggaagcca tgccgcattg gaggagaggt 300
 ttaccacaac ctgaagaatg tcatcaagga gaaatatggg aaagatgcc acaatgtggg 360
 gatttgcgog ggtttgctcc caacatcctg gagaataaag aaggcctgga gctgctgaag 420
 actgctattg gaaagcctgg cctacactgt aaagggtggc atggcatgga cgtagcggcc 480
 tccgagttct tcaggtcagg gaactatgac ctggacttca agtctcccga tgaccccagc 540
 aggtacatct cgctgacca gctggctgac ctgtacaagt ccttcatcaa ggactacca 600
 gtggtgtcta tcgaagatcc ctttgaccag gatgactggg gagcttcaga agttcacagc 660
 cagtgcagga atccaggtag tggggggatg actcacagt accaacccaa agaggatcgc 720
 caaggcgtga acgagaagtc ctgcaactgc ctctgtctca aagtcaacca gattggctcc 780
 gtgaccoggt ctcttcaggc gtgcaagctg gcccaggcca atggttgggg cgtcatggtg 840
 tctcatogtt cgggggagac tgaagatacc ttcatcgctg acctggttgt ggggctgtgc 900
 actggggcag atcaagactg gtgccccttg ccgatcacgc gcttggccaa gtacaaccag 960

```

ctcctcagaa ttgaagagga gctgggcagc aaggctaagt ttgccggcag gaacttcaga 1020
aacccttgg ccaagtaagc tgtgggcagg caagccttcg gtcacctgtt ggctacagac 1080
ccctcccttg gtgtcagctc aggcagctcg agggcccccga ccaacacttg caggggtccc 1140
tgctagttag cgccccccgc cgtggagttc gtaccgcttc cttagaactc tacagaagcc 1200
aagetccctg gaagccctgt tggcagctct agctttgcag ttgtgtaatt ggcccaagtc 1260
attgtttttc tcgccttact ttccaccaag tgtctagagt catgtgagcc tngtgtcatc 1320
tccggggtgg ccacaggcta gatccccggt ggttttgtgc tcaaaataaa aagcctcagt 1380
gaccatgaa aaaaaaaaaa gaattccgga attccg 1416

```

```

<210> 32
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> M55914

```

```

<400> 32
gtaccgcttc cttagaactc tacagaagcc aagctccctg gaagccctgt tggcagctct 60

```

```

<210> 33
<211> 2517
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> M96577

```

```

<400> 33
ggaattccgt ggccgggact ttgcaggcag cggcgcccg ggccgggagcg ggatcgagcc 60
ctgcgcgagg cctgcccga tggggccgcg ccgcgcgcgc cgctgtcac ccgggccgcg 120
cgggcccgtga gcgtcatggc cttggccggg gccctgctcg gcggcccatg cgcgcgcgcg 180
ctggaggccc tgctcggggc cggcgcgctg cggtgtctcg actcctcgca gatcgatc 240
atctccgcgc cgcaggacgc cagcgccccg ccggctccca ccggccccgc gccgcccgc 300
gccggccctt gcgacctga cctgctgctc ttgccacac cgcaggcgcc ccggcccaca 360
cccagtgcgc cgcggccgcg gctcggccgc ccgcgggtga agcggaggct ggacctggaa 420
actgaccatc agtacctggc cgagagcagt gggccagctc ggggcagagg ccgccatcca 480
ggaaaagggt tgaaatcccc gggggagaag tcacgctatg agacctact gaatctgacc 540
accaagcgct tcctggagct gctgagccac tcggctgacg gtgtcgtcga cctgaactgg 600
gctgccgagg tgctgaaggt gcagaagcgg cgcattctat acatcaccaa cgtccttgag 660
ggcatccagc tcattgccaa gaagtccaag aaccacatcc agtggctggg cagccacacc 720
acagtgggcg tcggcggacg gcttgagggg ttgaccagc acctccgaca gctgcaggag 780
agcgagcagc agctggacca cctgatgaat atctgtacta cgcagctgcg cctgctctcc 840
gaggacactg acagccagcg cctggcctac gtgacgtgtc aggaccttcg tagcattgca 900
gacctgagc agcagatggt tatggtgatc aaagccctc ctgagacca gctccaagcc 960
gtggactctt cggagaactt tcagatctcc cttaaagagca aacaaggccc gatcgatgtt 1020
ttcctgtgcc ctgaggagac cgtaggtggg atcagccctg ggaagacccc atcccaggag 1080
gtcacttctg aggaggagaa cagggccact gactctgcca ccatagtgtc accaccacca 1140
tcattctccc cctcatccct caccacagat ccagccagc ctctactcag cctggagcaa 1200
gaaccgctgt tgtcccgat gggcagccgt cgggctcccg tggacgagga ccgctgtcc 1260
cgctgggtgg cggccgactc gctcctggag catgtgcggg aggacttctc cggcctcctc 1320
cctgaggagt tcattagcct ttccccaccc cagcaggccc tcgactacca cttcggcctc 1380
gaggaggcg agggcatcag agactcttc gactgtgact ttggggaccc caccocctg 1440
gatttctgac agggcttggg gggaccaggg tttccagagt agctcacctt gtctctgag 1500
ccctggagcc ccctgtccct ggccgtctc ccagcctgtt tggaaacatt taatttatac 1560
ccctctctc tgtctccaga agcttctagc tctgggtct ggctaccgct aggaggctga 1620
gcaagccagg aagggaagga gtctgtgtg ttgtgtatgt catgcagcct acaccacac 1680
gtgtgtaccg ggggtgaatg tgtgtgagca tgtgtgtgt catgtaccg ggaatgaagg 1740
tgaacatata cctctgtgtg tgcactgcag acacgcccc gtgtgtccac atgtgtgtgc 1800
atgagcccat ctctgcgcgt gggggggctc taactgcact ttccggcctt ttgctcgtgg 1860

```

```

ggtccacaa ggcccagggc agtgcctgct cccagaatct ggtgctotga ccaggccagg 1920
tggggaggct ttggctggct gggcggtgag gacggtgaga gcacttctgt cttaaagggt 1980
ttttctgatt gaagctttaa tggagcgtaa tttatttatc gaggcctctt tggtagcct 2040
ggggaatcag caaaagggga ggaggggtgt ggggttgata ccccaactcc ctctaccctt 2100
gagcaagggc aggggtccct gagctgttct tctgccccat actgaaggaa ctgaggcctg 2160
ggtgatttat ttattgggaa agtgaggag ggagacagac tgactgacag ccatgggtgg 2220
tcagatggtg ggggtggggc tctccagggg gccagttcag ggcccagctg cccccaggga 2280

```

```

tgatatgag atgggagagg tgagtggggg accttactg atgtgggcag gaggggtggt 2340
gaaggcctcc cccagcccag accctgtggt ccttctgca gtgtctgaag cgcctgcctc 2400
cccactgctc tgccccaccc tccaatctgc actttgattt gcttcctaac agctctgttc 2460
cctctgctt tggttttaat aaatatattg atgacgttaa aaaaaggaat tcgatat 2517

```

<210> 34
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> M96577

<400> 34
 gtaggacggt gagagcactt ctgtcttaaa ggttttttct gattgaagct ttaatggagc 60

<210> 35
 <211> 4437
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000057

<400> 35
 gcgcggcgcc cgtgggtgcg gcgcgggaag tttggatcct gggtccgtcc gctaggagtc 60
 tgcggtcgag gattatggct gctgttcctc aaaataatct acaggagcaa ctagaacgtc 120
 actcagccag aacacttaat aataaattaa gtctttcaaa accaaaattt tcagggtttca 180
 cttttaaaaa gaaaacatct tcagataaca atgtatctgt aactaatgtg tcagtagcaa 240
 aaacacctgt attaagaaat aaagatgtta atgttaccga agacttttcc ttcagtgaac 300
 ctctacccaa caccacaaat cagcaaaggg tcaaggactt ctttaaaaaat gctccagcag 360
 gacaggaaac acagagaggt ggatcaaaat cattattgcc agatttcttg cagactccga 420
 aggaagtgt atgcactacc caaaacacac caactgtaaa gaaatcccg gatactgctc 480
 tcaagaaatt agaatttagt tcttcaccag attctttaag taccatcaat gattgggatg 540
 atatggatga ctttgatact tctgagactt caaaatcatt tgttacacca ccccaaagtc 600
 actttgtaag agtaagcact gctcagaaat caaaaaaggg taagagaaac ttttttaaa 660
 cacagcttta tacaacaaac acagtaaaga ctgatttgcc tccaccctcc tctgaaagcg 720
 agcaaataga tttgactgag gaacagaagg atgactcaga atgggtaagc agcgatgtga 780
 tttgcatcga tgatggcccc attgctgaag tgcataataa tgaagatgct caggaaagt 840
 actctctgaa aactcatttg gaagatgaaa gagataatag cgaaaagaag aagaatttgg 900
 aagaagctga attacattca actgagaaag ttccatgtat tgaatttgat gatgatgatt 960
 atgatacgga ttttgttcca ccttctccag aagaaattat ttctgcttct tcttctctct 1020
 caaaatgcct tagtacgtta aaggaccttg acacatctga cagaaaagag gatgttctta 1080
 gcacatcaaa agatcttttg tcaaaacctg agaaaatgag tatgcaggag ctgaatccag 1140
 aaaccagcac agactgtgac gctagacaga taagtttaca gcagcagctt attcatgtga 1200
 tggagacat ctgtaaaatta attgatacta ttctgtgata taaactgaaa cttttggatt 1260
 gtgggaacga actgcttcag cagcgggaaca taagaaggaa acttctaacg gaagtagatt 1320
 ttaataaaaag tgatgccagt cttcttggct cattgtggag atacaggcct gattcacttg 1380
 atggccctat ggagggtgat tcttgcctca cagggaattc tatgaaggag ttaaattttt 1440
 cacaccttcc ctcaaatctt gtttctcctg gggactgttt actgactacc accctaggaa 1500
 agacaggatt ctctgccacc aggaagaatc tttttgaaag gcctttattc aataccatt 1560
 tacagaagtc ctttgtaagt agcaactggg ctgaaacacc aagactagga aaaaaaatg 1620
 aaagctctta tttcccagga aatgttctca caagcactgc tgtgaaagat cagaataaac 1680

```

atactgcttc aataaatgac ttagaaagag aaacccaacc ttcctatgat attgataatt 1740
ttgacataga tgactttgat gatgatgatg actgggaaga cataatgcat aatttagcag 1800
ccagcaaatc ttccacagct gcctatcaac ccatcaagga aggtcggcca attaaatcag 1860
tatcagaaag actttcctca gccaaagacag actgtcttcc agtgtcatct actgctcaaa 1920
atataaactt ctcagagtca attcagaatt atactgacaa gtcagcacia aatttagcat 1980
ccagaaatct gaaacatgag cgtttccaaa gtcttagttt tcctcataca aaggaaatga 2040
tgaagatttt tcataaaaaa tttggcctgc ataattttag aactaatcag cttagggcga 2100
tcaatgctgc actgcttggg gaagactgtt ttatcctgat gccgactgga ggtggtaaga 2160
gtttgtgtta ccagctccct gcctgtgttt ctcctggggt cactgttgtc atttctccct 2220
tgagatcact tatcgtagat caagtccaaa agctgacttc cttggatatt ccagctacat 2280
atctgacagg tgataaagact gactcagaag ctacaaatat ttacctccag ttatcaaaaa 2340
aagacccaat cataaaactt ctatatgtca ctccagaaaa gatctgtgca agtaacagac 2400
tcattttctac tctggagaat ctctatgaga ggaagctctt ggcaagctttt gttattgatg 2460
aagcacattg tgtcagtcag tggggacatg attttctgca agattacaaa agaatagaata 2520
tgcttcgcca gaagtttccct tctgttccgg tgatggctct tacggccaca gctaatacca 2580
gggtacagaa ggacatcctg actcagctga agattctcag acctcagggt tttagcatga 2640
gctttaacag acataatctg aaataactatg tattaccgaa aaagcctaaa aaggtggcat 2700
ttgattgcct agaattggatc agaaagcacc acccatatga ttcagggata atttactgcc 2760
tctccaggcg agaattgtgac accatggctg acacgttaca gagagatggg ctcgctgtctc 2820
ttgcttacca tgcctggcctc agtgattctg ccagagatga agtgagcagc aagtggatta 2880
atcaggatgg ctgtcagggtt atctgtgcta caattgcatt tgggaatggg attgacaaac 2940
cggacgtgcy atttgtgatt catgcatctc tccttaaatc tgtggagggt tactaccaag 3000
aatctggcag agctggaaga gatggggaaa tatctcactg cctgcttttc tatacctatc 3060
atgatgtgac cagactgaaa agacttataa tgatggaaaa agatggaaac catcatacaa 3120
gagaaactca cttcaataat ttgtatagca tgggtacatta ctgtgaaaat ataacggaat 3180
gcaggagaat acagcttttg gcctactttg gtgaaaatgg atttaatcct gatttttcta 3240
agaaacaccc agatgtttct tgtgataatt gctgtaaaac aaaggattat aaaacaagag 3300
atgtgactga cagatgtgaaa agtattgtaa gatttgttca agaacatagt tcatcacaag 3360
gaatgagaaa tataaaacat gtaggtcctt ctggaagatt tactatgaat atgctggctg 3420
acattttctt ggggagtaag agtgcaaaaa tccagtcagg tatatttggg aaaggatctg 3480
cttatttcag acacaatgcc gaaagacttt ttaaaaagct gatacttgac aagatttttg 3540
atgaagactt atatatcaat gccaatgacc aggcgatcgc ttatgtgatg ctcggaataa 3600
aagcccaaac tgtactaaat ggcaatttaa aggtagactt tatggaaaca gaaaattcca 3660
gcagtgtgaa aaaacaaaaa gcgttagtag caaaagtgtc tcagagggaa gagatggtta 3720
aaaaatgtct tggagaactt acagaagtct gcaaatctct ggggaaagt tttggtgtcc 3780
attacttcaa tatttttaat accgtcactc tcaagaagct tgcagaatct ttatcttctg 3840
atcctgaggt tttgcttcaa attgatgggt ttactgaaga caaactggaa aaatatggtg 3900
cggaagtgat ttcagtatta cagaaatact ctgaatggac atcgccagct gaagacagtt 3960
ccccagggat aagcctgtcc agcagcagag gccccggaag aagtgcgctc gaggagcttg 4020
acgaggaat acccgtatct tcccactact ttgcaagtaa aaccagaaat gaaaggaaga 4080
ggaaaaagat gccagcctcc caaaggtcta agaggagaaa aactgcttcc agtggttcca 4140
aggcaaagg ggggtctgcc acatgtagaa agatatcttc caaaacgaaa tctccagca 4200
tcattggatc cagttcagcc tcacatactt ctcaagcgac atcaggagcc aatagcaaat 4260
tggggattat ggctccaccg aagcctataa atagaccgtt tcttaagcct tcatatgcat 4320
tctcataaca accgaatctc aatgtacata gacctcttt cttgtttgtc agcatctgac 4380
catctgtgac tataaagctg ttattcttgt tataccaaaa aaaaaaaaaa aaaaaaa 4437

```

<210> 36
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000057

<400> 36
 taagccttca tatgcattct cataacaacc gaatctcaat gtacatagac cctctttctt 60

<210> 37
 <211> 2016
 <212> DNA

<213> Homo sapiens

<300>

<308> NM_000060

<400> 37

gccagctgga	gcgttttcgg	ggctgtaaag	ggagaatggc	gcatgcgcat	attcagggcg	60
gaaggcgcg	taagagcaga	tttgtggtct	gcattatgtc	tggagccaga	agtaagcttg	120
ctcttttcc	ctgcggctgt	tacgtggttg	ccctgggagc	ccacaccggg	gaggagagcg	180
tggctgacca	tcacgaggct	gaatattatg	tggctgccgt	gtatgagcat	ccatccatcc	240
tgagtctgaa	ccctctggct	ctcatcagcc	gccaaagggc	cttggagctc	atgaaccaga	300

accttgacat	ctatgaacag	caagtgatga	ctgcagccca	aaaggatgta	cagattatag	360
tgtttccaga	agatggcatt	catggattca	actttacaag	aacatccatt	tatccatttt	420
tggacttcat	gccgtctccc	caggtggtca	ggtggaaccc	atgcctggag	cctcaccgct	480
tcaatgacac	agaggtgctc	cagcgctga	gttgatggc	catcagggga	gatatgttct	540
tgggtggcaa	tcttgggaca	aaggagcctt	gtcatagcag	tgaccaagg	tgcccaaaag	600
atgggaagata	ccagttcaac	acaaatgtcg	tgttcagcaa	taatggaacc	cttgttgacc	660
gctaccgtaa	acacaacctc	tactttgagg	cagcattcga	tgttcctctt	aaagtggatc	720
tcacacacct	tgataccccc	tttgcctggc	ggtttggcat	cttcacatgc	tttgatata	780
tgttctttga	ccctgccatc	agagtcctca	gagactacaa	ggtgaagcat	gttgtgtacc	840
caactgcctg	gatgaaccag	ctcccactct	tggcagcaat	tgagattcag	aaagcttttg	900
ctgttgccct	tggcatcaac	gttctggcag	ctaattgtcca	ccaccaggtt	ctggggatga	960
caggaaagtgg	catacacacc	cctctggagt	ccttttggtg	ccatgacatg	gaaaaatcca	1020
aaagtacact	tataattgcc	caggtggcca	aaaatccagt	gggtctcatt	ggtgcagaga	1080
atgcaacagg	tgaacaggac	ccatcccata	gtaagttttt	aaaaattttg	tcaggcgatc	1140
cgtactgtga	gaaggatgct	cagggaagtcc	actgtgatga	ggccaccaag	tggaacgtga	1200
atgctcctcc	cacatttcac	tctgagatga	tgtatgacaa	tttcaccctg	gtccctgtct	1260
ggggaaagga	aggctatctc	cacgtctgtt	ccaatggcct	ctgctgttat	ttactttacg	1320
agaggcccac	cttatccaaa	gagctgtatg	ccctgggggt	ctttgatggg	cttcacacag	1380
tacatggcac	ttactacatc	caagtgtgtg	ccctggctcag	gtgtgggggt	cttggcttcg	1440
acacctgcgg	acaggaatc	acagaggcca	cggggatatt	tgagtttcac	ctgtggggca	1500
acttcagtac	ttcctatata	tttcctttgt	ttctgacctc	agggatgacc	ctagaagatc	1560
ctgaccagct	tggctgggag	aatgaccact	atttcctgag	gaaaagtagg	ctgtcctctg	1620
ggctggtgac	ggcggctctc	tatgggcgct	tgtatgagag	ggactaggaa	aagtgtgtgg	1680
tctgtggggc	ggactctggc	catcatgttg	acagccttgc	acttccacag	gctacaagcc	1740
ctgggaccat	ctttctgcct	taagggcagg	agccccacttc	tgtggcacca	gattccaccc	1800
tgggaactgt	ggaaaaagta	ggagaggcag	attccctcag	tgtcttctct	ttaaacctca	1860
atcatcgaga	catttagggg	tattttctgt	tcacatttat	ctttttcaag	ccacatcttc	1920
ctctaacaaa	tctctcagta	tgcgattggt	ctcaagctaa	aacaaaaata	aatgtcagtt	1980
tatatatttac	acatccaaaa	aaaaaaaaaa	aaaaaa	2016		

<210> 38

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000060

<400> 38

tcctctaaca	aatctctcag	tatgcgattg	gtctcaagct	aaaacaaaaa	taaagtgcag	60
------------	------------	------------	------------	------------	------------	----

<210> 39

<211> 811

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000269

<400> 39

```

gcagaagcgt tccgtgcgtg caagtgcgtc gaaccacgtg ggtcccgggc gcgtttcggg 60
tgctggcggc tgcagccgga gttcaaacct aagcagctgg aaggaacat ggccaactgt 120
gagcgtaact tcattgcgat caaacagat ggggtccagc ggggtcttgt gggagagatt 180
atcaagcgtt ttgagcagaa aggattccgc cttgttggtc tgaaattcat gcaagcttcc 240
gaagatcttc tcaaggaaca ctacgttgac ctgaaggacc gtccattctt tgccggcctg 300
gtgaaataca tgcactcagg gccggtagt t gccatggtct gggaggggct gaatgtggtg 360
aagacggggc gagtcatgct cggggagacc aaccctgcag actccaagcc tgggaccatc 420
cgtggagact tctgcataca agttggcagg aacattatac atggcagtga ttctgtggag 480
agtgcagaga aggagatcgg cttgtggtt caccctgagg aactggtaga ttacacgagc 540
tgtgtcaga actggatcta tgaatgacag gagggcagac cacattgctt ttcacatcca 600
tttccctcc ttccatggg cagaggacca ggctgtagga aatctagtta ttacaggaa 660
cttcatcata atttgaggg aagctcttgg agctgtgagt tctccctgta cagtgttacc 720
atccccgacc atctgattaa aatgcttctt cccagcatag gattcattga gttggttact 780
tcatattgtt gcattgcttt ttttccctt t 811

```

<210> 40

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000269

<400> 40

```

gtctgaaatt catgcaagct tccgaagatc ttctcaagga acactacgtt gacctgaagg 60

```

<210> 41

<211> 2338

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000291

<400> 41

```

agcgcacgtc ggcagtcggc tccctcggtt accgaatcac cgacctctct cccagctgt 60
atttccaaaa tgcgctttc taacaagctg acgctggaca agctggacgt taaagggaag 120
cggttcgtta tgagatcga cttcaatgtt cctatgaaga acaaccagat aacaaacaac 180
cagagagatta aggtcgtgtt cccaagcatc aaattctgct tggacaatgg agccaagtgc 240
gtagtcctta tgagccacct aggcgggctt gatggtgtgc ccatgcctga caagtactcc 300
ttagagccag ttgctgtaga actcaaatct ctgctgggca aggatgttct gttcttgaag 360
gactgtgtag gcccagaagt ggagaaagcc tgtgccaacc cagctgctgg gtctgtcatc 420
ctgctggaga acctccgctt tcatgtggag gaagaaggga agggaaaaga tgcttctggg 480
aacaaggtta aagccgagcc agccaaaata gaagctttcc gagcttctact ttccaagcta 540
ggggtatgtc atgtcaatga tgcttttggc actgctcaca gagccacag ctccatggta 600
ggagtcaatc tgccacagaa ggctgggtgg tttttgatga agaaggagct gaactacttt 660
gcaaaggcct tggagagccc agagcgaccc ttctgggcca tcctggggcg agctaaagtt 720
gcagacaaga tccagctcat caataatatg ctggacaaag tcaatgagat gattattggt 780
ggtggaatgg cttttacctt ccttaagggt ctcaacaaca tggagattgg cacttctctg 840
tttgatgaag agggagccaa gattgtcaaa gacctaatgt ccaagctga gaagaatggt 900
gtgaagatta ccttgccgtg tgactttgtc actgctgaca agtttgatga gaatgccaa 960
actggccaag ccactgtggc ttctggcata cctgctggc ggatgggctt ggactgtggt 1020
cctgaaaaga gcaagaagta tgctgaggc gtcactcggg ctaagcagat tgtgtggaat 1080
ggctcctgtg gggatattga atgggaagct tttgcccggg gaaccaaagc tctcatggat 1140
gaggtggtga aagccacttc taggggctgc atcaccatca taggtggtgg agacactgcc 1200
acttgctgtg ccaaatggaa cacggaggat aaagtcagcc atgtgagcac tgggggtggt 1260
gccagtttgg agctcctgga aggtaaagtc ctctcctggg tggatgctct cagcaatatt 1320
tagtactttc ctgcctttta gttcctgtgc acagccccta agtcaactta gcattttctg 1380
catctccact tggcattagc taaaaccttc catgtcaaga ttcagctagt ggccaagaga 1440
tgacgtgcca ggaaccttta aacagttgca cagcatctca gctcatcttc actgcacct 1500

```

```

ggatttgcac acattcttca agatcccatt tgaatttttt agtgactaaa ccattgtgca 1560
ttctagagtg catatattta ttttttgccg gttaaaaaga aagtgagcag tgtagctta 1620
gttctctttt gatgtagggt attatgatta gctttgtcac tgtttacta ctcagcatgg 1680
aaacaagatg aaattccatt tgtaggtagt gagacaaaat tgatgatcca ttaagtaaac 1740
aataaaagtg tccattgaaa ccgtgatttt tttttttttc ctgtcatact ttgttaggaa 1800
gggtgagaat agaattctga ggaacggatc agatgtctat attgctgaat gcaagaagtg 1860
gggcagcagc agtggagaga tgggacaatt agataaatgt ccattcttta tcaaggccct 1920

```

```

actttatggc agacattgtg ctagtgcctt tattctaact tttattttta tcagttacac 1980
atgatcataa tttaaaaagt caaggcctat aacaaaaaag cccagccca ttcctcccat 2040
tcaagattcc cactcccagc aggtgaccac tttcaactct tgagtttttc aggtatatac 2100
ctccatgttt ctaagtaata tgcttatatt gtccacttcc ttttttttta ttttttaaag 2160
aaatctatct cataccatgg aggaaggctc tgttccacat atatttcac ttcctcattc 2220
tctcggtata gttttgtcac aattatagat tagatcaaaa gtctacataa ctaatacagc 2280
tgagctatgt agtatgctat gattaaatct acttatgtaa aaaaaaaaaa aaaaaaaa 2338

```

<210> 42
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000291

<400> 42
acttagcatt ttctgcatct ccacttggca ttagctaaaa ccttccatgt caagattcag 60

<210> 43
<211> 787
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000363

<400> 43
ctgaaggcca cccgggcggc cccctcactg accctccaaa cgccctgtc ctcgccctgc 60
ctcttgccat tcccggcctg agtctcagca tggcggtatg gagcagcagc gcggctaggg 120
aacctcgccc tgcaccagcc ccaatcagac gccgtcctc caactaccgc gcttatgcca 180
cggagcgcga cgccaagaaa aaatctaaga tctccgcctc gagaaaattg cagctgaaga 240
ctctgctgct gcagattgca aagcaagagc tggagcgaga gccggaggag cggcgaggag 300
agaagggcgc cgctctgagc acccgctgcc agccgctgga gttgaccggg ctgggcttcg 360
cggagctgca ggacttgtgc cgacagctcc acgccgtgt ggacaagggt gatgaagaga 420
gatacgacat agaggcaaaa gtcaccaaga acatcacgga gattgcagat ctgactcaga 480
agatctttga ccttcgaggc aagtttaagc ggcccaccct gcggagagtg aggatctctg 540
cagatgccat gatgcaggcg ctgctggggg cccgggctaa ggagtccttg gacctgcggg 600
cccacaccaa gcaggatgaag aaggaggaca ccgagaagga aaaccgggag gtgggagact 660
ggcggaagaa catcgatgca ctgagtggaa tggagggccg caagaaaaag tttgagagct 720
gagccttcct gcctactgcc cctgccctga ggagggccac tgaggaataa agcttctctc 780
tgagctg 787

<210> 44
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000363

<400> 44
tgtggacaag gtggatgaag agagatacga catagaggca aaagtcacca agaacatcac 60

PCT/US06/04280

<210> 45
 <211> 1263
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000365

<400> 45
 ggcacgagac cttcagcgcc tcggctccag cgccatggcg ccctccagga agttcttcgt 60
 tgggggaaac tggaagatga acggggcgaa gcagagtctg ggggagctca tcggcactct 120
 gaacgcggcc aaggtgccgg ccgacaccga ggtggtttgt gctcccccta ctgcctatat 180
 cgacttcgcc cggcagaagc tagatcccaa gattgctgtg gctgcgcaga actgctacaa 240
 agtgactaat ggggctttta ctggggagat cagccctggc atgatcaaag actgcgagc 300
 cacgtgggtg gtcctggggc actcagagag aaggcatgtc tttggggagt cagatgagct 360
 gattgggcag aaagtggccc atgctctggc agagggactc ggagtaatcg cctgcattgg 420
 ggagaagcta gatgaaaggg aagctggcat cactgagaag gttgttttcg agcagacaaa 480
 ggtcatcgca gataacgtga aggactggag caaggtcgtc ctggcctatg agcctgtgtg 540
 ggccattggt actggcaaga ctgcaacacc ccaacaggcc caggaagtac acgagaagct 600
 ccgaggatgg ctgaagtcca acgtctctga tgcggtggct cagagcaccg gtatcattta 660
 tggaggtctc gtgactgggg caacctgcaa ggagctggcc agccagcctg atgtggatgg 720
 cttccttgtg ggtggtgctt ccctcaagcc cgaattcgtg gacatcatca atgccaaaca 780
 atgagcccca tccatcttcc ctacccttcc tgccaagcca gggactaagc agcccagaag 840
 ccagtaacta gccctttccc tgcatactgt tctgatgggt tcatctgctc cttcctgtgg 900
 cctcatccaa actgtatctt cctttactgt ttatatcttc accctgtaat ggttgggacc 960
 aggccaatcc cttctccact tactataatg gttggaacta aacgtcacca aggtggcttc 1020
 tccttggtcg agagatggaa ggcgtggtgg gatttgctcc tgggttcctt aggccttagt 1080
 gagggcagaa gagaaaccat cctctccctt cttacaccgt gaggccaaga tcccctcaga 1140
 aggcaggagt gctgccctct cccatgggtg ccgtgcctct gtgctgtgta tgtgaaccac 1200
 ccatgtgagg gaataaacct ggcactagga aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1260
 aaa 1263

<210> 46
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000365

<400> 46
 tatcttcacc ctgtaatggt tgggaccagg ccaatccctt ctccacttac tataatggtt 60

<210> 47
 <211> 1616
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000582

<400> 47
 ctccctgtgt tggtaggagga tgtctgcagc agcattttaa ttctgggagg gcttggttgt 60
 cagcagcagc aggaggaggc agagcacagc atcgctcgga ccagactcgt ctccaggccag 120
 ttgcagcctt ctacgcaaaa cgccgaccaa ggaaaactca ctaccatgag aattgcagtg 180
 atttgctttt gcctcctagg catcacctgt gccataccag ttaaacaggc tgattctgga 240
 agttctgagg aaaagcagct ttacaacaaa taccagatg ctgtggccac atggctaaac 300
 cctgacccat ctgagaagca gaatctccta gcccacaga cccttccaag taagtccaac 360
 gaaagccatg accacatgga tgatatggat gatgaagatg atgatgacca tgtggacagc 420
 caggactcca ttgactcgaa cgactctgat gatgtagatg aactgatga ttctcaccag 480

```

tctgatgagt ctcaccattc tgatgaatct gatgaactgg tcaactgattt tcccacggac 540
ctgccagcaa ccgaagtgtt cactccagtt gtccccacag tagacacata tgatggccga 600
ggtgatagtg tggtttatgg actgagggtca aaatctaaga agtttcgcag acctgacatc 660
cagtaccctg atgctacaga cgaggacatc acctcacaca tggaaagcga ggagttgaat 720
ggtgcataca aggccatccc cgttgcccag gacctgaacg cgccttctga ttgggacagc 780
cgtgggaagg acagtatatga aacgagtcag ctggatgacc agagtgtctga aaccacacagc 840
cacaagcagt ccagattata taagcggaaa gccaatgatg agagcaatga gcattccgat 900
gtgattgata gtcaggaact ttccaaagtc agccgtgaat tccacagcca tgaatttcac 960
agccatgaag atatgctggt tgtagacccc aaaagtaagg aagaagataa acacctgaaa 1020
tttcgtatgt ctcattgaatt agatagtgc tcttctgagg tcaattaaaa ggagaaaaaa 1080
tacaatttct cactttgcat ttagtcaaaa gaaaaaatgc tttatagcaa aatgaaagag 1140
aacatgaaat gcttctttct cagtttattg gttgaatgtg tatctatttg agtctggaaa 1200
taactaatgt gtttgataat tagtttagtt tgtggcttca tggaaactcc ctgtaaacta 1260
aaagcttcag ggttatgtct atgttcattc tatagaagaa atgcaaacta tcaactgtatt 1320
ttaatatattg ttattctctc atgaatagaa atttatgtag aagcaaacta aatactttta 1380
cccacttaa aagagaatat aacattttat gtcactataa tcttttggtt ttaagttag 1440
tgtatatatt gttgtgataa tcttttggtg gtgtgaataa atcttttatc ttgaatgtaa 1500
taagaatttg gtgggtgtcaa ttgcttattt gttttccac ggttgtccag caattaataa 1560
aacataacct tttttactgc ctaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 1616

```

<210> 48
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000582

<400> 48
 ggtggtgtca attgcttatt tgttttccca cggttgtcca gcaattaata aaacataacc 60

<210> 49
 <211> 1666
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000584

<400> 49
 ctccataagg cacaaacttt cagagacagc agagcacaca agcttctagg acaagagcca 60
 ggaagaaacc accggaagga accatctcac tgtgtgtaaa catgacttcc aagctggccg 120
 tggtctctct ggcagccttc ctgatttctg cagctctgtg tgaaggtgca gttttgccaa 180
 ggagtgtctaa agaacttaga tgtcagtgc taaagacata ctccaaacct ttccacccca 240
 aatttatcaa agaactgaga gtgattgaga gtggaccaca ctgcgccaac acagaaatta 300
 ttgtaaagct ttctgatgga agagagctct gtctggaccc caaggaaaac tgggtgcaga 360
 gggttgtgga gaagtttttg aagagggtct agaattcata aaaaaattca ttctctgtgg 420
 tatccaagaa tcagtgaaga tgccagtga aactcaagca aatctacttc aacacttcat 480
 gtattgtgtg ggtctgttgt agggttgcca gatgcaatac aagattcctg gttaaatttg 540
 aatttcagta aacaatgaat agtttttcat tgtaccatga aatatccaga acatacttat 600
 atgtaaagta ttattttatt gaatctacaa aaaacaacaa ataattttta aatataagga 660
 ttttctaga tattgcacgg gagaatatac aaatagcaaa attgaggcca agggccaaga 720
 gaatatccga actttaattt caggaattga atgggtttgc tagaatgtga tatttgaagc 780
 atcacataaa aatgatggga caataaattt tgccataaag tcaaatttag ctggaaatcc 840
 tggatttttt tctgttaaat ctggcaacct tagtctgcta gccaggatcc acaagtcctt 900
 gtccactgt gccttggttt ctcttttatt tctaagtgga aaaagtatta gccaccatct 960
 tacctcacag tgatgttgtg aggacatgtg gaagcacttt aagttttttc atcataacat 1020
 aaattatttt caagtgtaac ttattaacct atttattatt tatgtattta ttttaagcatc 1080
 aaatatttgt gcaagaattt ggaaaaatag aagatgaatc attgattgaa tagttataaa 1140
 gatgttatag taaattttatt ttattttaga tattaaatga tgttttatta gataaatttc 1200
 aatcagggtt ttttagattaa acaaacaac aattgggtac ccagttaaat tttcatttca 1260

```

gataaacaac aaataatttt ttagtataag tacattattg tttatctgaa attttaattg 1320
aactaacaat cctagtttga tactcccagt cttgtcattg ccagctgtgt tggtagtgct 1380
gtgttgaatt acggaataat gagttagaac tattaataca gccaaaactc cacagtcaat 1440
attagtaatt tcttgctggt tgaaacttgt ttattatgta caaatagatt cttataatat 1500
tattttaaag actgcatttt taaatacaag gcttttatatt tttaacttta agatgttttt 1560
atgtgctctc caaatttttt ttactgtttc tgattgtatg gaaatataaa agtaaatatg 1620
aaacatttaa aatataattt gttgtcaaag taataaaaaa aaaaaa 1666

```

<210> 50

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000584

<400> 50

```

tggtagtgct gtgttgaatt acggaataat gagttagaac tattaataca gccaaaactc 60

```

<210> 51

<211> 1722

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000599

<400> 51

```

ggggaaaaga gctaggaaaag agctgcaaag cagtgtgggc tttttccctt tttttgctcc 60
ttttcattac cctcctccg ttttcaccct tctccggact tcgcgtagaa cctgcgaatt 120
tcgaagagga ggtggcaaaag tgggagaaaa gaggtgtagg ggtttggggt ttttttggtt 180
ttgtttttgt tttttaattt cttgatttca acattttctc ccacctctc ggtgcagcc 240
aacgctctt acctgttctg cggcgccgcg caccgctggc agctgagggg tagaaagcgg 300
ggtgtatttt agattttaag caaaaatttt aaagataaat ccatttttct ctcccacccc 360
caacgccatc tccactgcat ccgatctcat tatttcgggtg gttgcttggg ggtgaacaat 420
tttggtgctt tttttccctt ataattctga cccgctcagg cttgaggggt tctccggcct 480
ccgctcactg cgtgcacctg gcgctgcctt gcttccccc acctgttgca aggctttaat 540
tcttgcaact gggacctgct cgcaggcacc ccagccctcc acctctctc acatttttgc 600
aagtgtctgg gggagggcac ctgctctacc tgccagaaat tttaaaacaa aaacaaaaac 660
aaaaaaatct cggggggccc tcttgcccc tttatccctg cactctcgct ctctgcccc 720
accccgaggt aaagggggcg actaagagaa gatggtgttg ctacccgagg tcctcctgct 780
gctggccgcc tatgcccggc cggcccagag cctgggctcc ttcgtgcact gcgagccctg 840
cgacgagaaa gccctctcca tgtgcccccc cagcccccct ggctgcgagc tggtaagga 900
gccgggctgc ggctgctgca tgacctgcgc cctggccgag gggcagtcgt gcggcgtcta 960
caccgagcgc tgcgccagg ggctgcgctg cctcccccgg caggacgagg agaagccgct 1020
gcacgccctg ctgcacggcc ggggggtttg cctcaacgaa aagagctacc gcgagcaagt 1080
caagatcgag agagactccc gtgagcacga ggagcccacc acctctgaga tggccgagga 1140
gacctactcc cccaagatct tccggcccaa acacaccgc atctccgagc tgaaggctga 1200
agcagtgaag aaggaccgca gaaagaagct gaccagctcc aagtttgtcg ggggagccga 1260
gaacactgcc cacccccgga tcatctctgc acctgagatg agacaggagt ctgagcaggg 1320
cccctgcgcg agacacatgg aggttccct gcaggagctc aaagccagcc cacgcattgt 1380
gcccgtgtgt gtgtacctgc ccaattgtga ccgcaaaagg ttctacaaga gaaagcagt 1440
caaaccttcc cgtggccgca agcgtggcat ctgctggtgc gtggacaagt acgggatgaa 1500
gctgccagga atggagtaag ttgacgggga ctttcagtgc cacaccttcg acagcagcaa 1560
cgttgagtga tgcgtcccc cccaaccttt ccctcaccct ctcccacccc cagccccgac 1620
tccagccagc gcctccctcc accccaggac gccactcatt tcatctcatt taagggaata 1680
atatatatct atctatttga ggaaaaaaa aaaaaaaa aa 1722

```

<210> 52

<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_000599

<400> 52
ccaggagcgc actcatttca tctcatttaa gggaaaaata tataatctatc tatttgagga 60

<210> 53
<211> 704
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000735

<400> 53
gcagttactg agaactcata agacgaagct aaaatccctc ttcggatcca cagtcaaccg 60
ccctgaacac atcctgcaaa aagcccagag aaaggagcgc catggattac tacagaaaat 120
atgcagctat ctttctggtc acattgtcgc tgtttctgca tgttctccat tccgctcctg 180
atgtgcagga ttgccagaa tgacgctac aggaaaaccc attcttctcc cagccgggtg 240
ccccaaactc tcagtgcatt ggctgctgct tctctagagc atatccact ccactaaggt 300
ccaagaagac gatgttggtc caaagaacg tcacctcaga gtccacttgc tgtgtagcta 360
aatcatataa cagggtcaca gtaatggggg gtttcaaagt ggagaaccac acggcgtgcc 420
actgcagtac ttgttattat cacaaatctt aaatgtttta ccaagtgtg tcttgatgac 480
tgctgatttt ctggaatgga aaattaagtt gtttagtggt tatggctttg tgagataaaa 540
ctctcctttt ccttaccata ccactttgac acgcttcaag gatatactgc agctttactg 600
ccttcctcct tatectacag tacaatcagc agtctagttc ttttcatttg gaatgaatac 660
agcattaagc ttgttccact gcaaataaag ccttttaaat catc 704

<210> 54
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000735

<400> 54
tgagataaaa ctctcctttt ccttaccata ccactttgac acgcttcaag gatatactgc 60

<210> 55
<211> 1342
<212> DNA
<213> Homo sapiens

<300>
<308> NM_000799

<400> 55
cccgagccg gaccggggcc accgcgccg ctctgctccg acaccggcc ccctggacag 60
ccgccctctc ctccaggccc gtggggctgg ccctgcacgc ccgagcttcc cgggatgagg 120
gccccgggtg tgggtacccg gcgcgcccc ggtcgtgag ggacccggc caggcgcgga 180
gatgggggtg cacgaatgtc ctgcctggct gtggcttctc ctgtccctgc tgtcgtctcc 240
tctgggcctc ccagtccctg gcgcccacc acgcctcatc tgtgacagcc gagtccctga 300
gaggtacctc ttggaggcca aggaggccga gaatatcagc acgggctgtg ctgaacactg 360
cagcttgaat gagaatatca ctgtcccaga caccaaagt aatttctatg cctggaagag 420
gatggaggtc gggcagcagg ccgtagaagt ctggcagggc ctggccctgc tgtcgggaagc 480
tgtcctgcgg ggccaggccc tgttggtcaa ctcttccag ccgtgggagc cctgcagct 540
gcatgtgat aaagccgtca gtggccttcg cagcctcacc actctgctc gggctctgcg 600

```

agcccagaag gaagccatct cccctccaga tgcggcctca gctgctccac tccgaacaat 660
cactgtgtac actttccgca aactcttccg agtctactcc aatttccctc ggggaaagct 720
gaagctgtac acaggggagg cctgcaggac aggggacaga tgaccagggtg tgtccacctg 780
ggcatatcca ccacctccct caccaacatt gcttgtgcca caccctcccc cgccactcct 840
gaaccccgtc gaggggctct cagctcagcg ccagcctgtc ccatggacac tccagtgcc 900
gcaatgacat ctcaggggcc agaggaactg tccagagagc aactctgaga tctaaggatg 960
tcacagggcc aacttgaggg cccagagcag gaagcattca gagagcagct ttaaactcag 1020
ggacagagcc atgctgggaa gacgcctgag ctcactcggc accctgcaaa atttgatgcc 1080
aggacacgct ttggaggcga ttacctgtt ttcgcacctt ccatcaggga caggatgacc 1140
tgagaaactt aggtggcaag ctgtgacttc tccaggtctc acgggcatgg gcactccctt 1200
ggtggcaaga gcccccttga caccggggtg gtgggaacca tgaagacagg atgggggctg 1260
gcctctggct ctcattgggt ccaagttttg tgtattcttc aacctcattg acaagaactg 1320
aaaccaccaa aaaaaaaaaa aa 1342

```

<210> 56
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000799

```

<400> 56
tcattggggtc caagttttgt gtattcttca acctcattga caagaactga aaccacaaaa 60

```

<210> 57
 <211> 2722
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000917

```

<400> 57
gagcgggctg agggtaggaa gtagccgctc cgagtggagg cgactggggg ctgaagagcg 60
cgccgccctc tcgtcccact ttccagggtg gtgacccctg aaaattaaat cttccaagat 120
gatctggtat atattaatta taggaattct gcttccccag tctttggctc atccaggctt 180
ttttacttca attggtcaga tgactgattt gatccatact gagaaagatc tggtgacttc 240
tctgaaagat tatattaagg cagaagagga caagttagaa caaataaaaa aatgggcaga 300
gaagttgat cggctaacta gtacagcgac aaaagatcca gaaggatttg ttgggcatcc 360
agtaaatgca ttcaaattaa tgaaacgtct gaatactgag tggagtgagt tggagaatct 420
ggtccttaag gatatgtcag atggctttat ctctaacctt accattcaga gaccagtact 480
ttctaagat gaagatcagg ttggggcagc caaagctctg ttacgtctcc aggataccta 540
caatttgat acagatacca tctcaaaggg taatcttcca ggagtgaac acaaatcttt 600
tctaacggct gaggactgct ttgagttggg caaagtggcc tatacagaag cagattatta 660
ccatacgga ctgtggatgg aacaagccct aaggcaactg gatgaaggcg agatttctac 720
catagataaa gtctctgttc tagattattt gagctatgcg gtatatcagc agggagacct 780
ggataaggca cttttgtcga caaagaagct tcttgaacta gatcctgaac atcagagagc 840
taatggtaac ttaaaatatt ttgagtatat aatggctaaa gaaaaagatg tcaataagtc 900
tgcttcagat gaccaatctg atcagaaaac tacaccaaag aaaaaagggg ttgctgtgga 960
ttacctgcc gagagacaga agtacgaaat gctgtgccgt ggggagggtg tcaaatgac 1020
ccctcgaga cagaaaaaac tcttttgccg ataccatgat ggaaaccgta atcctaaatt 1080
tattctggct ccagctaaac aggaggatga ctgggacaag cctcgatta ttgccttoca 1140
tgatattatt tctgatgcag aaattgaaat cgtcaaagac ctacgaaaac caaggctgag 1200
ccgagctaca gtacatgacc ctgagactgg aaaattgacc acagcacagt acagagtatc 1260
taagagtgcc tggctctctg gctatgaaaa tctgtgggtg tctogaatta atatgagaat 1320
acaagatcta acaggactag atgtttccac agcagaggaa ttacaggtag caaattatgg 1380
agttggagga cagtatgaac cccattttga ctttgacagg aaagatgagc cagatgcttt 1440
caaagagctg gggacaggaa atagaattgc tacatggctg ttttatatga gtgatgtgtc 1500
tgagagga gccactgtt ttctgaagt tggagctagt gtttggcca aaaaaggaac 1560
tgctgttttc tggataatc tgtttgccag tggagaagga gattatagta cacggcatgg 1620

```



```

agcctgtcca gtgctagttg gcaacaaatg ggtatccaat aaatggctcc atgaacgtgg 1680
acaagaatth cgaagacctt gtacgttggtc agaattggaa tgacaaacag gcttcccttt 1740
ttctcttatt gttgtactct tatgtgtctg atatacacat ttccatagtc ttaactttca 1800
ggagttttaca attgactaac actccatgat tgattcagtc atgaacctca tcccatgttt 1860
catctgtgga caattgctta ctttgtgggt tcttttaaaa gtaacacgaa atcatcata 1920
tgcataaaac cttaaagtgc tggtgggtatc acagaagaca aggcagagtt taaagtgagg 1980
aatthttatat ttaaagaact ttttgggttg ataaaaacat aatttgagca tccagtttta 2040
gtattttcact acatctcagt tgggtgggtgt taagctagaa tgggctgtgt gataggaaac 2100
aaatgcctta cagatgtgcc taggtgttct gtttacctag tgtcttactc tgttttctgg 2160
atctgaagac tagtaataaa ctaggacact aactgggttc catgtgattg ccctttcata 2220
tgatcttcta agttgatttt tttctcccca agtctttttt aaagaaagta tactgtattt 2280
taccaccccc ctctcttttc ttttagctcc tctgtgggtg attaaacgta cttgagttaa 2340
aatatttoga tttttttttt ttttttaatg gaaagtcctg cataacaaca ctgggccttc 2400
ttaactaaaa tgctcaccac ttagcctgtt tttttatccc ttttttaaaa tgacagatga 2460
ttttgttcag gaattttgtt gtttttctta gtgctaatac cttgcctctt attcctgcta 2520
cagcagggtg gtaatatgtg cattctgatt aaatactgtg ccttaggaga ctggaagttt 2580
aaaaaatgtac aagtcctttc agtgatgagg gaattgattt tttttaaaag tcttttctt 2640
agaaagccaa aatgtttgtt tttttaagat tctgaaatgt gttgtgacaa caatgaccta 2700
tttatgatct taaatctttt tt 2722

```

<210> 58

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000917

<400> 58

```

tcttactctg ttttctggat ctgaagacta gtaataaact aggacactaa ctgggttcca 60

```

<210> 59

<211> 3236

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001109

<400> 59

```

gaccgggcca tgcgcgccct cgggctctgg ctgctgggag cgatgatgct gcctgcgatt 60
gccccagacc ggccctgggc cctcatggag cagtatgagg tctgtgtgcc gcggcgctctg 120
ccaggccccc gagtccgccc agctctgccc tccacttggt gcctgcaccc agagaggggtg 180
agctacgtcc ttggggccac agggcacaaac ttcaccctcc acctgcggaa gaacagggag 240
ctgctgggtt ccggctacac agagacctat acggctgcca atggctccga ggtgacggag 300
cagcctcgcg ggccaggcca ctgcttatac cagggccacg tagaggggta ccgggactca 360
gcccgcagcc tcagcacctg tgccggcctc aggggtttct tccaggtggg gtcagacctg 420
cacctgatcg agccctgga tgaaggtggc gagggcggac ggacgcctgt gtaccaggct 480
gagcacctgc tgcagacggc cgggacctgc ggggtcagcg acgacagcct gggcagcctc 540
ctgggacccc ggacggcagc cgtcttcagg cctcgccccc gggactctct gccatcccga 600
gagaccogct acgtggagct gtatgtggtc gtggacaatg cagagttcca gatgctgggg 660
agcgaagcag ccgtgcgtca tccgggtgct gaggtgggtg atcacgtgga caagctatat 720
cagaaactca acttccgtgt ggtcctgggt ggctggaga tttggaatg tcaggacagg 780
ttccacgtca gcccgcagcc cagtgtcaca ctggagaacc tctgacctg gcaggcagcg 840
caacggacac ggccggcacct gcatgacaac gtacagctca tcacgggtgt cgaacttcacc 900
gggactactg tggggtttgc cagggtgtcc gccatgtgct cccacagctc aggggtgtgtg 960
aaccaggacc acagcaagaa ccccggtggc gtggcctgca ccatggccca tgagatgggc 1020
cacaacctgg gcatggacca tgatgagaac gtccagggtc gccgctgcca ggaacgcttc 1080
gaggccggcc gctgcacat ggcaggcagc attggctcca gtttccccag gatgttcagt 1140
gactgcagcc aggcctacct ggagagcttt ttggagcggc cgcagtcggt gtgcctcgcc 1200
aacgccctcg acctcagcca cctgggtggc ggccccgtgt gtgggaacct gtttgtggag 1260

```

```

cgtgggggagc agtgcgactg cggccccccc gaggactgcc ggaaccgctg ctgcaactct 1320
accacctgcc agctggctga gggggcccag tgtgcgcacg gtacctgctg ccaggagtgc 1380
aaggtgaagc cggctggtga gctgtgccgt cccaagaagg acatgtgtga cctcgaggag 1440
ttctgtgacg gccgcacccc tgagtgcccg gaagacgcct tccaggagaa cggcacgccc 1500
tgctccgggg gctactgcta caacggggcc tgtcccacac tggcccagca gtgccaggcc 1560
ttctgggggc caggtgggca ggctgccgag gagtccctgt tctcctatga catcctacca 1620
ggctgcaagg ccagccggtg cagggctgac atgtgtggcg ttctgcagtg caagggtggg 1680
cagcagcccc tggggcgctg catctgcata gtggatgtgt gccacgcgct caccacagag 1740
gatggcactg cgtatgaacc agtgcccag ggcacccggt gtggaccaga gaaggtttgc 1800
tggaaggac gttgccagga cttacacgtt tacagatcca gcaactgctc tgcccagtgc 1860
cacaaccatg ggggtgtgcaa ccacaagcag gagtgccact gccacgcggg ctgggccccg 1920
ccccactgcg cgaagtgtgt gactgagggt cagcagcgct ccgggagcct ccccgctctc 1980
gtggtggtgg ttctggtgct cctggcagtt gtgctggtca ccctggcagg catcatcgctc 2040
taccgcaaag cccggagccg catcctgagc aggaacgtgg ctccaagac cacaatgggg 2100
cgctccaacc ccctgttcca ccaggctgcc agccgcgtgc cggccaaggg cggggctcca 2160
gccccatcca ggggccccca agagctggtc ccaccacccc acccgggcca gcccgccgca 2220
cacccggcct cctcggtggc tctgaagagg ccgccccctg ctctccggt cactgtgtcc 2280
agcccacct tcccagttcc tgtctacacc cggcaggcac caaagcaggt catcaagcca 2340
acgttcgcac cccagtgcc cccagtcaaa cccggggctg gtgcggccaa ccctggtcca 2400
gctgaggggt ctgttggccc aaaggttgcc ctgaagcccc ccatccagag gaagcaagga 2460
gccggagctc ccacagcacc ctaggggggc acctgcgcct gtgtggaaat ttggagaagt 2520
tgccgagagc aagccatgcg ttccagcctt ccacggtcca gctagtgccg ctacgcccta 2580
gacctgact ttgcaggtc agctgctgtt ctaacctcag taatgcatct acctgagagg 2640
ctcctgctgt ccacgccctc agccaattcc ttctccccgc cttggccacg tgtagcccca 2700
gctgtctgca ggcaccaggc tgggatgagc tgtgtgcttg cgggtgcgtg tgtgtgtacg 2760
tgtctccagg tggccgctgg tctcccgctg tgttcaggag gccacatata cagcccctcc 2820
cagccacacc tgcccctgct ctggggcctg ctgagccggc tgcccctggc acccggttcc 2880
aggcagcaca gacgtggggc atccccagaa agactccatc ccaggaccag gttcccctcc 2940
gtgtctctcg agagggtgct agtgagcaga ctgcacccca agctcccagc tccaggctcc 3000
ctgatcttgg gcctgtttcc catgggattc aagagggaca gcccagctt tgtgtgtgtt 3060
taagcttagg aatgcccttt atggaaaggg ctatgtggga gagtacagta tcttgtctgg 3120
tttcttgag acctcagatg tgtgttcagc agggctgaaa gctttttattc ttttaataatg 3180
agaaatgtat attttactaa taaattattg accgagttct gtagattctt gttaga 3236

```

<210> 60

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001109

<400> 60

ctttatggaa agggctatgt gggagagtca gctatcttgt ctggttttct tgagacctca 60

<210> 61

<211> 1449

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001124

<400> 61

```

ctggatagaa cagctcaagc cttgccactt cgggcttctc actgcagctg ggcttggact 60
tcggagtttt gccattgcca gtgggacgtc tgagactttc tcottcaagt acttggcaga 120
tcactctctt agcagggtct gcgcttcgca gccgggatga agctggtttc cgtcgccctg 180
atgtacctgg gttcgctcgc cttcctaggc gctgacaccg ctcggttgga tgtcgcgctc 240
gagtttcgaa agaagtggaa taagtgggct ctgagtcgtg ggaagaggga actgcggatg 300
tccagcagct accccaccgg gctcgctgac gtgaaggccg ggccctgcca gacccttatt 360
cggccccagg acatgaaggg tgccctctcga agccccgaag acagcagtcg ggatgcgcgc 420

```

```

cgcatccgag tcaagcgota ccgccagagc atgaacaact tccagggcct ccggagcttt 480
ggctgccgct tcgggacgtg cacggtgcag aagctggcac accagatcta ccagttcaca 540
gataaggaca aggacaacgt cgcccccagg agcaagatca gcccccagg ctacggccgc 600
cggcgcgggc gtcctctgcc cgaggccggc ccgggtcgga ctctggtgtc ttctaagcca 660
caagcacacg gggctccagc ccccccgagt ggaagtgtct cccactttct ttaggattta 720
ggcgcccatg gtacaaggaa tagtcgcgca agcatcccgc tgggtgcctcc cgggacgaag 780
gacttcccga gcggtgtggg gaccgggctc tgacagccct gcggagacct tgagtccggg 840
aggcacccgtc cggcggcgag ctctggcttt gcaagggccc ctccctcttg gggcttcgct 900
tccttagcct tgctcaggtg caagtgtccc agggggcggg gtgcagaaga atccgagtgt 960
ttgccaggct taaggagagg agaaactgag aaatgaatgc tgagaccccc ggagcagggg 1020
tctgagccac agccgtgtct gccacaaac tgatttctca cggcgtgtca cccaccagg 1080
gcgcaagcct cactattact tgaactttcc aaaacctaaa gaggaaaagt gcaatgcgtg 1140
ttgtacatac agaggttaact atcaatattt aagtttgttg ctgtcaagat tttttttgta 1200
acttcaaata tagagatatt tttgtacgtt atatattgta ttaagggcat tttaaaagca 1260
attatattgt cctccctat ttttaagacgt gaatgtctca gcgaggtgta aagttgttcg 1320
ccgctgtgaa tgtgagtgtg tttgtgtgca tgaaagagaa agactgatta cctcctgtgt 1380
ggaagaagga aacaccgagt ctctgtataa tctatttaca taaaatgggt gatatgcgaa 1440
cagcaaac 1449

```

<210> 62

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001124

<400> 62

```

gaaggaaaca ccgagtctct gtataatcta tttacataaa atgggtgata tgccaacagc 60

```

<210> 63

<211> 1619

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001168

<400> 63

```

ccgccagatt tgaatcgcg gacccgttgg cagaggtggc ggccggcgga tgggtgcccc 60
gacgttgccc cctgcctggc agccctttct caaggaccac cgcattctcta cattcaagaa 120
ctggcccttc ttggagggtc gcgcctgcac cccggagcgg atggccgagg ctggcttcac 180
ccactgcccc actgagaacg agccagactt ggcccagtggt ttcttctgct tcaaggagct 240
ggaaggctgg gagccagatg acgaccccat agaggaacat aaaaagcatt cgtccgggtg 300
cgctttcctt tctgtcaaga agcagtttga agaattaacc cttggtgaat ttttgaaact 360
ggacagagaa agagccaaga acaaaattgc aaaggaaacc aacaataaga agaaagaatt 420
tgaggaaaact gcgaagaaag tgcgccgtgc catcgagcag ctggctgcca tggattgagg 480
cctctggccg gagctgcctg gtcccagagt ggctgcacca cttccagggt ttattccctg 540
gtgccaccag cttcctgtg ggccccttag caatgtctta ggaaaggaga tcaacatttt 600
caaattagat gtttcaactg tgctcctgtt ttgtcttgaa agtggcacca gaggtgcttc 660
tgctgtgca gcgggtgtct ctggttaacag tggctgcttc tctctctctc tctctttttt 720
gggggtcat ttttgcgtgt ttgattcccg ggcttaccag gtgagaagtg agggaggaag 780
aaggcagtggt cccttttgc agagctgaca gctttgttcg cgtgggcaga gccttcacca 840
gtgaatgtgt ctggacctca tgttgttgag gctgtcacag tcttgagtgt ggacttggca 900
ggtgcctgtt gaatctgagc tgcaggttcc ttatctgtca cacctgtgcc tccctagagg 960
acagtttttt tgttgtgtgt tttttttgtt tttttttttt ggtagatgca tgacttgtgt 1020
gtgatgagag aatggagaca gagtccctgg ctcccttact gtttaacaac atggctttct 1080
tattttgttt gaattgttaa ttcacagaat agcacaaact acaattaaaa ctaagcacia 1140
agccattcta agtcattggg gaaacggggt gaacttcagg tggatgagga gacagaatag 1200
agtgatagga agcgtctggc agatactcct tttgccactg ctgtgtgatt agacaggccc 1260
agtgaagcgc gggggcacatg ctggccgctc ctccctcaga aaaaggcagt ggccataatc 1320

```

```

ctttttaaat gacttggtc gatgctgtgg gggactggct gggctgctgc aggcctgtg 1380
tctgtcagcc caaccttcac atctgtcacg ttctccacac gggggagaga cgcagtccgc 1440
ccaggctccc gctttctttg gaggcagcag ctcccgcagg gctgaagtct ggcgtaagat 1500
gatggatttg attcgccctc ctccctgtca tagagctgca ggggtgattg ttacagcttc 1560
gctggaaacc tctggaggtc atctcggtcg ttcttgagaa ataaaaagcc tgtcatttc 1619

```

```

<210> 64
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_001168

```

```

<400> 64
ttcacagaat agcacaaact acaattaaaa ctaagcacia agccattcta agtcattggg 60

```

```

<210> 65
<211> 1552
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_001216

```

```

<400> 65
gcccgtacac accgtgtgct gggacacccc acagtcagcc gcatggctcc cctgtgcccc 60
agcccctggc tccctctgtt gatcccgccc cctgctccag gcctcactgt gcaactgctg 120
ctgtcactgc tgcttctgat gcctgtccat ccccagaggt tgccccggat gcaggaggat 180
tcccccttgg gaggaggctc ttctggggaa gatgacccac tgggcgagga ggcctgtccc 240
agtgaagagg attcaccacag agaggaggat ccacccggag aggaggatct acctggagag 300
gagatctac ctggagagga ggatctacct gaagttaagc ctaaatacaga agaagagggc 360
tccctgaagt tagaggatct acctactgtt gaggctcctg gagatcctca agaaccacag 420
aataatgccc acagggacaa agaaggggat gaccagagtc attggcgcta tggaggcgac 480
ccgcccctgg cccgggtgtc cccagcctgc gggggcctct tccagtcccc ggtggatata 540
cgccccagc tcgcccctt ctgcccggcc ctgcccggcc tggaaactct gggcttccag 600
ctcccggcgc tcccagaact gcgcctgcgc aacaatggcc acagtgtgca actgaccctg 660
cctcctgggc tagagatggc tctgggtccc gggcgaggat accgggctct gcagctgcat 720
ctgactggg gggctgcagg tcgtccgggc tcggagcaca ctgtggaagg ccaccgtttc 780
cctgccgaga tccacgtggt tcacctcagc accgctttg ccagagttga cgaggcctt 840
gggcgcccgg gaggcctggc cgtgttggcc gcctttcttg aggagggcc ggaagaaaac 900
agtgcctatg agcagttgct gtctcgcttg gaagaaatcg ctgaggaagg ctgagagact 960
caggtcccag gactggacat atctgcaact ctgcccctct acttcagccg ctacttccaa 1020
tatgaggggt ctctgactac accgcctgt gccagggtg tcatctggac tgtgtttaac 1080
cagacagtga tgctgagtgc taagcagctc cacaccctct ctgacaccct gtggggacct 1140
ggtgactctc ggctacagct gaacttccga gcgacgcagc ctttgaatgg gcgagtgatt 1200
gaggcctcct tccctgctgg agtgagacag agtcctcggg ctgctgagcc agtccagctg 1260
aatctctgcc tggctgctgg tgacatccta gccctggttt ttggcctcct ttttgetgtc 1320
accagectcg cgttccttgt gcagatgaga aggcagcaca gaaggggaac caaagggggt 1380
gtgagctacc gccagcaga ggtagccgag actggagcct agaggctgga tcttgagaaa 1440
tgtgagaagc cagccagagg catctgagg ggagccggta actgtcctgt cctgctcatt 1500
atgccacttc cttttaactg ccaagaaatt ttttaaaata aatatttata at 1552

```

```

<210> 66
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_001216

```

<400> 66

tcctgtcctg ctcattatgc cacttccttt taactgccaa gaaatttttt aaaataaata 60

<210> 67

<211> 2653

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001254

<400> 67

gagcgcggtt ggagtttgct gctgccgctg tgcagtttgt tcaggggctt gtggtggtga 60
 gtccgagagg ctgctgtgta gagacgtgag aaggatcctg cactgaggag gtggaaagaa 120
 gaggattgct cgaggaggcc tggggtctgt gaggcagcgg agctgggtga aggctgagg 180
 ttccggcgag gcctgagctg tgctgtcgtc atgcctcaaa cccgatccca ggcacaggct 240
 acaatcagtt ttccaaaaag gaagctgtct cgggcattga acaaagctaa aaactccagt 300
 gatgccaaac tagaaccac aaatgtccaa accgtaacct gttctcctcg tgtaaaagcc 360
 ctgctctca gccccaggaa acgtctgggc gatgacaacc tatgcaacac tccccattta 420
 cctccttgtt ctccaccaa gcaaggcaag aaagagaatg gtccccctca ctcacatata 480
 cttaagggac gaagattggt atttgacaat cagctgacaa ttaagtctcc tagcaaaaga 540
 gaactagcca aagttacca aaacaaaata ctttcttcag ttagaaaaag tcaagagatc 600
 acaacaaatt ctgagcagag atgtccactg aagaaagaat ctgcatgtgt gagactattc 660
 aagcaagaag gcacttgcta ccagcaagca aagctggctc tgaacacagc tgtcccagat 720
 cggctgctgt ccagggaag ggagatggat gtcacagga atttcttgag ggaacacatc 780
 tgtgggaaaa aagctgggaag cctttacctt tctgggtgctc ctggaaactgg aaaaactgcc 840
 tgcttaagcc ggattctgca agacctcaag aaggaactga aaggctttaa aactatcatg 900
 ctgaattgca tgccttgtag gactgcccag gctgtattcc cagctattgc tcaggagatt 960
 tgtcaggaag aggtatccag gccagctggg aaggacatga tgaggaaatt ggaaaaacat 1020
 atgactgcag agaaggggcc catgattgtg ttggtattgg acgagatgga tcaactggac 1080
 agcaaggcc aggatgtatt gtacacgcta tttgaatggc catgggctaag caattctcag 1140
 ttggtgctga ttggtattgc taataccctg gatctcacag atagaattct acctaggctt 1200
 caagctagag aaaaatgtaa gccacagctg ttgaacttcc cacttatac cagaaatcag 1260
 atagtcacta ttttgcaaga tcgacttaat caggatctca gagatcaggt tctggacaat 1320
 gctgcagttc aattctgtgc ccgcaaagtc tctgctgttt caggagatgt tcgcaaagca 1380
 ctggatgttt gcaggagagc tattgaaatt gtagagtcag atgtcaaaaag ccagactatt 1440
 ctcaaaccac tgtctgaatg taaatcacct tctgagcctc tgattcccaa gagggttggt 1500
 ctattcaca tatcccaagt catctcagaa gttgatggta acaggatgac cttgagccaa 1560
 gaaggagcac aagattcctt cctcttcag cagaagatct tggtttgctc tttgatgctc 1620
 ttgatcaggc agttgaaat caaagaggct actctgggga agttatatga agcctacgt 1680
 aaagtctgtc gcaaacagca ggtggcggct gtggaccagt cagagtgttt gtcaactttca 1740
 gggctcttgg aagccagggg cattttagga ttaaagagaa acaaggaaac ccgtttgaca 1800
 aagggtgttt tcaagattga agagaaagaa atagaacatg ctctgaaaga taaagcttta 1860
 attgaaaata tcttagctac tggattgcct taaattcttc tcttacacc caccgaaag 1920

tattcagctg gcatttagag agctacagtc ttcatcttag tgctttacac attcgggctt 1980
 gaaaacaaat atgacctttt ttacttgaag ccaatgaatt ttaattctata gattctttta 2040
 tattagcaca gaataatata tttgggtctt actattttta cccataaaaag tgaccaggta 2100
 gaccttttt aattacattc actacttcta ccacttggtg atctctagcc aatgtgcttg 2160
 caagtgtaca gatctgtgta gaggaatgtg tgtatatatta cctcttcgtt tgctcaaaca 2220
 tgagtgggta tttttttgtt tgtttttttt gttgtgtgtt tttttgagge gcgtctcacc 2280
 ctggtgcccc ggctggagtg caatggcgcg ttctctgctc actacagcac ccgcttccca 2340
 ggttgaagtg attctcttgc ctgagcctcc cgagtatgtg ggattacagg tgccccacc 2400
 cgcgccagc taatttttta attttttagta gagacagggt ttaccatgt tggccaggct 2460
 ggtcttgaac tctgacctt caagtgatct gccaccttg gcctccctaa gtgctgggat 2520
 tataggcgtg agccaccatg ctgagccatt aagggtatctt gttaagaact ttaagtttag 2580
 ggtaagaaga atgaaaatga tocagaaaaa tgcaagcaag tccacatgga gatttgagg 2640
 acactggtta aag 2653

<210> 68

<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_001254

<400> 68
caaggaaacc cgtttgacaa aggtgttttt caagattgaa gagaaagaaa tagaacatgc 60

<210> 69
<211> 627
<212> DNA
<213> Homo sapiens

<300>
<308> NM_001323

<400> 69
gcgccgcaa gctcggcact cacggctctg agggctccga cggcactgac ggccatggcg 60
cgttcgaacc tcccgtggc gctgggcctg gccctggtcg cattctgcct cctggcgctg 120
ccacgcgacg cccgggcccg gccgcaggag cgcattgctg gagaactccg ggacctgtcg 180
cccgcgcacc cgcaggtgca gaaggcggcg caggcggccg tggccagcta caacatgggc 240
agcaacagca tctactactt ccgagacacg cacatcatca aggcgcagag ccagctgggtg 300
gccggcatca agtacttctt gacgatggag atggggagca cagactgccg caagaccagg 360
gtcactggag accacgtcga cctcaccact tgccccctgg cagcaggggc gcagcaggag 420
aagctgcgct gtgactttga ggtccttggt gttccctggc agaactcctc tcagctccta 480
aagcacaact gtgtgcagat gtgataagtc cccgagggcg aaggccattg gggttggggc 540
catggtggag ggcacttcag gtccgtgggc cgtatctgtc acaataaatg gccagtgtctg 600
cttcttgcaa aaaaaaaaaa aaaaaaa 627

<210> 70
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_001323

<400> 70
atcaagtact tcctgacgat ggagatgggg agcacagact gccgcaagac cagggtcact 60

<210> 71
<211> 1812
<212> DNA
<213> Homo sapiens

<300>
<308> NM_001428

<400> 71
tagctaggca ggaagtccgc gcgggcggcg cggacagtat ctgtgggtac ccggagcacg 60
gagatctcgc cggctttacg ttcacctcgg tgtctgcagc accctccgct tcctctccta 120
ggcgacgaga cccagtggct agaagttcac catgtctatt ctcaagatcc atgccaggga 180
gatctttgac tctcgcggga atcccactgt tgaggttgat ctcttcacct caaaaggtct 240
cttcagagct gctgtgcccc gtggtgcttc aactggatc tatgaggccc tagagctccg 300
ggacaatgat aagactcgct atatggggaa ggggtgtctc aaggctgttg agcacatcaa 360
taaaactatt gcgcctgccc tggtagcaa gaaactgaac gtcacagaac aagagaagat 420
tgacaaactg atgatcgaga tggatggaa agaaaataaa tctaagtttg gtgcgaacgc 480
cattctgggg gtgtcccttg ccgtctgcaa agctgggtgc gttgagaagg gggccccct 540
gtaccgccac atcgctgact tggctggcaa ctctgaagtc atcctgccag tcccggcggt 600
caatgtcatc aatggcgggt ctcagtctgg caacaagctg gccatgcagg agttcatgat 660

```

cctcccagtc ggtgcagcaa acttcagggg agccatgcgc attggagcag aggtttacca 720
caacctgaag aatgtcatca aggagaaata tgggaaagat gccaccaatg tgggggatga 780
aggcggggttt gctcccaaca tcctggagaa taaagaaggc ctggagctgc tgaagactgc 840
tattgggaaa gctggctaca ctgataagggt ggatcatcggc atggacgtag cggcctccga 900
gttcttcagg tctgggaagt atgacctgga cttcaagtct cccgatgacc ccagcaggta 960
catctgcctt gaccagctgg ctgacctgta caagtccttc atcaaggact acccagtggg 1020
gtctatcgaa gatccctttg accaggatga ctggggagct tggcagaagt tcacagccag 1080
tgcaggaatc caggtagtgg gggatgatct cacagtgacc aacccaaaga ggatcgccaa 1140
ggcgtgaac gagaagtcct gcaactgcct cctgctcaaa gtcaaccaga ttggctccgt 1200
gaccgagtct cttcaggcgt gcaagctggc ccaggccaat ggttggggcg tcatggtgtc 1260
tcatcgttcg ggggagactg aagatacctt catectgac ctggttgtgg ggctgtgcac 1320
tgggcagatc aagactggtg ccccttgccg atctgagcgc ttggccaagt acaaccagct 1380
cctcagaatt gaagaggagc tgggcagcaa ggctaagtt gccggcagga acttcagaaa 1440
ccccttgccc aagtaagctg tgggcaggca agcccttcgg tcacctgttg gctacacaga 1500
cccctccccct cgtgtcagct caggcagctc gaggcccccg accaactt gcagggtcc 1560
ctgctagtta gcgccccacc gccgtggagt tcgtaccgct tccttagaac ttctacagaa 1620
gccaagctcc ctggagccct gttggcagct ctagctttgc agtcgtgtaa ttggcccaag 1680
tcatgttttt tctgcctca ctttccacca agtgtctaga gtcagtgtag cctcgtgtca 1740
tctccggggg gccacaggc tagatccccg gtggttttgt gctcaaaata aaaagcctca 1800
gtgaccatg ag 1812

```

<210> 72

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001428

<400> 72

```

agctctagct tttgcagtcg tgtaatgggc ccaagtcatt gtttttctcg cctcactttc 60

```

<210> 73

<211> 8368

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001456

<400> 73

```

gcgatccggg cggcaccgcc cgggtcatcg tcaccgggtcg ctctcaggaa cagcagcgca 60
acctctgctc cctgcctcgc ctcccgcgcg cctaggtgcc tgcgacttta attaaagggc 120
cgtcccctcg ccgaggctgc agcaccgccc ccccggttc tcgcgcctca aaatgagtag 180
ctcccaactct cgggcgggcc agagcgcagc aggcgcggct ccgggcggcg gcgtcgacac 240
gcgggacgcc gagatgccgg ccaccgagaa ggacctggcg gaggacgcgc cgtggaagaa 300
gatccagcag aacactttca cgcgtggtg caacgagcac ctgaagtgcg tgagcaagcg 360
catcgccaac ctgcagacgg acctgagcga cgggctgcgc cttatcgcg tgttggaggt 420

```

```

gctcagccag aagaagatgc accgcaagca caaccagcgg cccactttcc gccaaatgca 480
gcttgagaac gtgtcgttg cgctcagatt cctggaccgc gagagcatca aactggtgtc 540
catcgacagc aaggccatcg tggacgggaa cctgaagctg atcctgggcc tcactctggac 600
cctgatcctg cactactcca tctccatgcc catgtgggac gaggaggagg atgaggaggc 660
caagaagcag accccaaagc agaggctcct gggctggatc cagaacaagc tgccgcagct 720
gcccataacc aacttcagcc gggactggca gagcggccgg gccctgggcg ccctggtgga 780
cagctgtgcc ccgggcctgt gtccctgactg ggactcttg gacgccagca agcccgttac 840
caatgcgcga gaggccatgc agcaggcgga tgactggctg ggcatcccc aggtgatcac 900
ccccgaggag attgtggacc ccaacgtgga cgagcactct gtcagtacct acctgtccca 960
gttccccaag gccaaagtga agccaggggc tcccttgccg cccaaactga acccgaagaa 1020
agcccgtgcc tacgggccag gcatcgagcc cacaggcaac atgggtgaaga agcgggcaga 1080
gttcaactgtg gagaccagaa gtgctggcca gggagagggt ctggtgtacg tggaggaccc 1140

```

ggccggacac	caggaggagg	caaaagtac	cgccaataac	gacaagaacc	gcaccttctc	1200
cgtctggtac	gtccccgagg	tgacggggac	tcataaggtt	actgtgctct	ttgctggcca	1260
gcacatcgcc	aagagccct	tcgaggtgta	cgtggataag	tcacagggtg	acgccagcaa	1320
agtacagcc	caaggtcccg	gcctggagcc	cagtggcaac	atcgccaaca	agaccaccta	1380
ctttgagatc	tttacggcag	gagctggcac	ggcgagggtc	gaggttgtga	tccaggaccc	1440
catgggacag	aagggcacgg	tagagcctca	gctggaggcc	cggggcgaca	gcacataccg	1500
ctgcagctac	cagcccacca	tggaggcggt	ccacaccgtg	cacgtcacgt	ttgccggcgt	1560
gcccacccct	cgagccccc	acactgtcac	tgttggtcaa	gcctgttaacc	cgagtgcctg	1620
ccggggcggt	ggccggggcc	tccagcccaa	gggtgtgcgg	gtgaaggaga	cagctgactt	1680
caagggtgtac	acaaagggcg	ctggcagtg	ggagctgaag	gtcaccgtga	agggcccca	1740
gggagaggag	cgctgaagc	agaaggacct	gggggatggc	gtgtatggct	tcgagtatta	1800
ccccatggtc	cctggaacct	atatcgtcac	catcacgtgg	ggtggtcaga	acatcgggcg	1860
cagtcccttc	gaagtgaagg	tgggcaccga	gtgtggcaat	cagaaggtag	gggcctgggg	1920
cctgggctg	gagggcgcg	tcgttgcaa	gtcagcagac	tttgtgggtg	aggctatcgg	1980
ggacgacgtg	ggcagcgtg	gcttctcggt	ggaaggccca	tcgaggcta	agatcgaatg	2040
tgacgacaag	ggcgacggct	cctgtgatgt	gcgctactgg	ccgaggagg	ctggcgagta	2100
tgccgttgca	gtgtgtgca	acaggaaga	catccgctc	agccccttca	tggctgacat	2160
ccctgacgcg	ccccagact	tccaccaga	cagggtgaag	gcacgtgggc	ctggattgga	2220
gaagacaggt	gtggcgtca	acaagccagc	agagttcaca	gtggatgcca	agcacggtgg	2280
caaggcccca	cttcgggtcc	aagtccagga	caatgaaggc	tgccctgtgg	aggcgttggt	2340
caaggacaac	ggcaatggca	cttacagctg	ctcctacgtg	cccaggaagc	cggtgaagca	2400
cacagccatg	gtgtcctggg	gaggcgtcag	catccccaac	agccccttca	gggtgaatgt	2460
gggagctggc	agccacccca	acaaggtcaa	agtatacggc	cccgagtag	ccaagacagg	2520
gctcaaggcc	cacgagccca	cctacttcac	tgtggactgc	gccgaggctg	gccaggggga	2580
cgtcagcatc	ggcatcaagt	gtgccctgg	agtggtaggc	cccgccgaag	ctgacatcga	2640
cttcgacatc	atccgcaatg	acaatgacac	cttcacggtc	aagtacacgc	cccggggggc	2700
tggcagctac	accattatgg	tcctctttgc	tgaccaggcc	acgccacca	gccccatccg	2760
agtcaagggtg	gagccctctc	atgacgccag	taagggtgaag	gccgagggcc	ctggcctcag	2820
tcgactgggt	gtcgagcttg	gcaagccac	ccacttcaca	gtaaatgcc	aagctgctgg	2880
caaaggcaag	ctggacgtcc	agttctcagg	actcaccaag	ggggatgcag	tgcgagatgt	2940
ggacatcatc	gaccaccatg	acaacaccta	cacagtcaag	tacacgcctg	tccagcaggg	3000
tcacgtaggc	gtcaatgtca	cttatggagg	ggatcccatc	cctaagagcc	ctttctcagt	3060
ggcagtatct	ccaagcctgg	acctcagcaa	gatcaagggt	tctggcctgg	gagagaaggt	3120
ggacgttggc	aaagaccagg	agttcacagt	caaatacaag	ggtgctgggtg	gtcaaggcaa	3180
agtggtcatc	aagattgtgg	gcccctcggg	tgcagcgggtg	ccctgcaagg	tggagccagg	3240
cctgggggct	gacaacagtg	tgggtgcgctt	cctgccccgt	gaggaagggc	cctatgaggt	3300
ggaggtgacc	tatgacggcg	tgcccggtcc	tggcagcccc	ttcctctctg	aagctgtggc	3360
ccccaccaag	cctagcaagg	tgaaggcgtt	tgggcccggg	ctgcaggagg	gcagtgcggg	3420
ctccccgcgc	cgcttcacca	tcgacaccaa	gggcgcgggc	acaggtggcc	tgggocctgac	3480
ggtggagggc	ccctgtgagg	cgcagctcga	gtgcttggtg	aatggggatg	gcacatgttc	3540
cgtgtcctac	gtgcccaccg	agcccgggga	ctacaacatc	aacatcctct	tcgctgacac	3600
ccacatccct	ggctcccat	tcaaggccca	cgtggttccc	tgctttgacg	catccaaagt	3660
caagtgtctca	ggccccgggc	tggagcgggc	caccgctggg	gaggtggggc	aattccaagt	3720
ggactgctcg	agcgcgggca	gcgcggagct	gaccattgag	atctgctcgg	aggcggggct	3780
tcgggcccag	gtgtacatcc	aggaccacgg	tgatggcacg	cacaccatta	cctacattcc	3840
cctctgcccc	ggggcctaca	ccgtcaccat	caagtagcgc	ggccagcccc	tggccaactt	3900
ccccgcaag	ctgcaggtgg	aacctgcggt	ggcacttccc	ggtgtccagt	gctatgggcc	3960
tggatttgag	ggccagggtg	tcttccgtga	ggccaccact	gagttcagtg	tggacggccc	4020
ggctctgaca	cagaccggag	ggccgcacgt	caaggcccgt	gtggccaacc	cctcaggcaa	4080
cctgacggag	acctacgttc	aggaccgtgg	cgtatggcatg	tacaaagtgg	agtacacgcc	4140
ttacgaggag	ggactgcact	ccgtggacgt	gacctatgac	ggcagtcctg	tggccagcag	4200
ccccctccag	gtgcccgtga	ccgagggtcg	cgacccctcc	cgggtgcgtg	tccacggggc	4260
aggcatccaa	agtggtgacca	ccaacaagcc	caacaagttc	actgtggaga	ccagggggagc	4320
tggcacgggc	ggcctgggcc	tggctgtaga	gggcccctcc	gaggccaaga	tgtcctgcat	4380
ggataacaag	gacggcagct	gctcgggtcga	gtacatccct	tatgaggctg	gcacctacag	4440
cctcaacgtc	acctatgggtg	gccatcaagt	gccaggcagt	cctttcaagg	tccctgtgca	4500
tgatgtgaca	gatgcgtcca	aggtcaagtg	ctctggggccc	ggcctgagcc	caggcatggt	4560
tcgtgccaac	ctccctcagt	ccttccaggt	ggacacaagc	aaggctgggtg	tggccccatt	4620
gcaggtcaaa	gtgcaagggc	ccaaaggcct	ggtggagcca	gtggacgtgg	tagacaacgc	4680
tgatggcacc	cagaccgtca	attatgtgcc	cagccgagaa	gggcccctaca	gcattctcagt	4740
actgtatgga	gatgaagagg	taccccgag	ccccttcaag	gtcaagggtgc	tgccctactca	4800

tgatgccagc	aaggtgaagg	ccagtggccc	cggtctcaac	accactggcg	tgcctgccag	4860
cctgcccgtg	gagttcacca	tcgatgcaaa	ggagcccg	gagggcctgc	tggctgtcca	4920
gatcacggat	cccgaaggca	agccgaagaa	gacacacac	caagacaacc	atgacggcac	4980
gtatacagtg	gcctacgtgc	cagacgtgac	aggtcgctac	accatcctca	tcaagtacgg	5040
tggtagcag	atcccccttct	ccccgtaccg	cgtgcgtgcc	gtgcccaccg	gggacgccag	5100
caagtgcact	gtcacagtgt	caatcggagg	tcacgggcta	ggtgctggca	tcggccccac	5160
cattcagatt	ggggaggaga	cggtgatcac	tgtggacact	aaggcggcag	gcaaaggcaa	5220
agtacgtgc	accgtgtgca	cgcctgatgg	ctcagaggtg	gatgtggacg	tggtagagaa	5280
tgaggacggc	actttcgaca	tcttctacac	ggccccccag	ccgggcaaat	acgtcatctg	5340
tgtgcgcttt	gtggcgagc	acgtgcccaa	cagccccttc	caagtgcagg	ctctggcttg	5400
ggacacggcc	tcggtgcagc	cccctctacg	gtctcagcag	ctggccccac	agtacaccta	5460
cgcccagggc	ggccagcaga	cttgggcccc	ggagaggccc	ctgggtgggtg	tcaatgggct	5520
ggatgtgacc	agcctgaggc	cctttgacct	tgtcatcccc	ttcaccatca	agaaggcgga	5580
gatcacaggg	gaggttcgga	tgcctcagg	caaggtggcg	cagcccacca	tactgacaa	5640
caaagacggc	accgtgaccg	tgcggtatgc	accagcgag	gctggcctgc	acgagatgga	5700
catccgctat	gacaacatgc	acatcccagg	aagccccttg	cagttctatg	tggattacgt	5760
caactgtggc	catgtcactg	cctatgggcc	tggcctcacc	catggagtag	tgaacaagcc	5820
tgccaccttc	accgtcaaca	ccaaggatgc	aggagagggg	ggcctgtctc	tggccattga	5880
ggggccgtcc	aaagcagaaa	tcagctgcac	tgacaaccag	gatgggacat	gcagcgtgtc	5940
ctacctgcct	gtgctgccgg	gggactacag	cattctagtc	aagtacaatg	aacagcacgt	6000
cccaggcagc	cccttactg	ctcgggtcac	aggtgacgac	tccatgcgta	tgtccacct	6060
aaaggtcggc	tctgctgccg	acatccccat	caacatctca	gagacggatc	tcagcctgct	6120
gacggccaact	gtggtcccgc	cctcgggccc	ggaggagccc	tgtttgctga	agcggctgcg	6180
taatggccac	gtggggattt	cattcgtgcc	caaggagacg	ggggagcacc	tggtagatgt	6240
gaagaaaaat	ggccagcacg	tggccagcag	ccccatcccg	gtggtgatca	gccagtcgga	6300
aattggggat	ccagctcgtg	tccgggtctc	tggtagggc	cttcacgaag	gccacacctt	6360
tgagcctgca	gagtttatca	ttgatacccg	cgtgcaggc	tatggtgggc	tcagcctgtc	6420
cattgagggc	cccagcaagg	tggacatcaa	cacagaggac	ctggaggagc	ggacgtgcag	6480
ggtcacctac	tgccccacag	agccaggcaa	ctacatcatc	aacatcaagt	ttgccgacca	6540
gcacgtgcct	ggcagcccct	tctctgtgaa	ggtgacaggc	gagggccggg	tgaagagag	6600
catcacccgc	aggcgtcggg	ctccttcagt	ggccaacgtt	ggtagtcatt	gtgacctcag	6660
cctgaaaatt	cctgaaaatt	gcattccagg	tatgacagcc	caggtgacca	gcccatcggg	6720
caagaccat	gaggccgaga	tcgtggaagg	tgagaaccac	acctactgca	tccgctttgt	6780
tcccgtgag	atgggacac	acacagtcag	cgtcaagtac	aagggccagc	acgtgcctgg	6840
gagccccttc	cagttcaccc	tggggcccct	aggggaagg	ggagcccaca	aggtccgagc	6900
tggggggcct	ggcctggaga	gagctgaagc	tggagtggca	gccgaattca	gtatctggac	6960
ccgggaagct	ggtgctggag	gcctggccat	tgtgtgcag	ggccccagca	aggctgagat	7020
ctcttttgag	gaccgcaagg	acggctcctg	tgggtgtggc	tatgtgtgtc	aggagccagg	7080
tgactacgaa	gtctcagtc	agttcaacga	ggaacacatt	cccgcagacc	ccttcgtggt	7140
gcctgtggct	tctccgtctg	gogacgccc	ccgcctcact	gtttctagcc	ttcaggagtc	7200
agggctaaag	gtcaaccagc	cagcctcttt	tgcagtcagc	ctgaacgggg	ccaagggggc	7260
gatcgatgcc	aaggtgcaca	gcccctcagg	agccctggag	gagtgctatg	tcacagaaat	7320
tgaccaagat	aagtatgctg	tgcgcttcat	ccctcgggag	aatggcgctt	acctgattga	7380
cgtcaagttc	aacggtaccc	acatcccctg	aagccccttc	aagatccgag	ttggggagcc	7440
tgggcatgga	ggggacccag	gcttgggtgc	tgtttacgga	gcaggtctgg	aaggcgggtg	7500
cacagggaac	ccagctgagt	tcgtcgtgaa	cacgagcaat	gcgggagctg	gtgccctgtc	7560
ggtgaccatt	gacggcccct	ccaaggtgaa	gatggattgc	caggagtgc	ctgagggcta	7620
ccgcgtcacc	tataccccca	tggcacctgg	cagctacctc	atctccatca	agtacggcgg	7680
cccctaccac	attgggggca	gccccttcaa	ggccaaagtc	acaggccccc	gtctcgtcag	7740
caaccacagc	ctccacgaga	catcatcagt	gtttgtagac	tctctgacca	aggccacctg	7800
tgccccccag	catggggccc	cgggtcctgg	gcctgctgac	gccagcaagg	tggtagccaa	7860
gggctgagg	ctgagcaagg	cctacgtagg	ccagaagagc	agcttcacag	tagactgcag	7920
caaagcaggc	aacaacatgc	tgtgtgtggg	ggttcatggc	ccaaggaccc	cctgcgagga	7980
gatcctgggtg	agcagcgtgg	gcagccggct	ctacagcgtg	tcctacctgc	tcaaggacaa	8040
gggggagtag	acactggtgg	tcaaatgggg	gcacgagcac	atcccaggca	gcccctaccg	8100
cgttgtgtgtg	ccctgagttc	ggggcccgtg	ccagccggca	gcccccaagc	ctgccccgct	8160
acccaagcag	ccccgccttc	ttcccctcaa	ccccggccca	ggccgccttg	gccgcccggc	8220
tgtcactgca	gctgcccctg	ccctgtgccg	tgtgtgcgtc	acctgcctec	ccagccagcc	8280
gctgacctct	cggctttcac	ttgggcagag	ggagccattt	ggtggcgctg	cttgtcttct	8340
ttggttctgg	gaggggtgag	ggatgggg	8368			

<210> 74
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001456

<400> 74
 tgacctctcg gctttcactt gggcagaggg agccatttgg tggcgctgct tgtcttcttt 60

<210> 75
 <211> 1642
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001548

<400> 75
 ccagatctca gaggagcctg gctaagcaaa accctgcaga acggctgcct aatttacagc 60
 aaccatgagt acaaatggtg atgatcatca ggtcaaggat agtctggagc aattgagatg 120
 tcactttaca tgggagttat ccattgatga cgatgaaatg cctgatttag aaaacagagt 180
 cttggatcag attgaattcc tagacaccaa atacagtgtg ggaatacaca acctactagc 240
 ctatgtgaaa cacctgaaag gccagaatga ggaagccctg aagagcttaa aagaagctga 300
 aaacttaatg caggaagaac atgacaacca agcaaattgtg aggagtctgg tgacctgggg 360
 caactttgcc tggatgtatt accacatggg cagactggca gaagcccaga cttacctgga 420
 caagggtggag aacatttgcga agaagctttc aaatcccttc cgctatagaa tggagtgtcc 480
 agaaatagac tgtgaggaag gatgggcctt gctgaagtgt ggaggaaaga attatgaacg 540
 ggccaaggcc tgctttgaaa aggtgcttga agtggaccct gaaaaccctg aatccagcgc 600
 tgggtatgcg atctctgcct atcgccctgga tggctttaaa ttagccacaa aaaatcacia 660
 gccattttct ttgcttcccc taaggcaggc tgtccgctta aatccagaca atggatatat 720
 taaggttctc cttgccctga agcttcagga tgaaggacag gaagctgaag gagaaaagta 780
 cattgaagaa gctctagcca acatgtcttc acagacctat gtctttcgat atgcagccaa 840
 gttttaccga agaaaaggct ctgtggataa agctcttgag ttattaaaaa aggccttgca 900
 ggaaacaccc acttctgtct tactgcatca ccagataggg ctttgctaca aggcacaaat 960
 gatccaaatc aaggaggcta caaaagggca gcctagaggg cagaacagag aaaagctaga 1020
 caaaatgata agatcagcca tatttcattt tgaatctgca gtggaaaaaa agcccacatt 1080
 tgagggtgct catctagacc tggcaagaat gtatatagaa gcaggcaatc acagaaaagc 1140
 tgaagagaat tttcaaaaat tgttatgcat gaaaccagtg gtagaagaaa caatgcaaga 1200
 catacatttc tactatggtc ggtttcagga atttcaaaaag aaatctgacg tcaatgcaat 1260
 tatccattat ttaaaagcta taaaaataga acaggcatca ttaacaaggg ataaaagtat 1320
 caattctttg aagaaattgg ttttaaggaa acttcggaga aaggcattag atctggaaaag 1380
 cttgagcctc cttgggttcg tctataaatt ggaaggaaat atgaatgaag ccctggagta 1440
 ctatgagcgg gccctgagac tggctgctga ctttgagaac tctgtgagac aaggtcctta 1500
 ggcaccacaga tatcagccac tttcacattt catttcattt tatgctaaca tttactaatc 1560
 atcttttctg cttactgttt tcagaaacat tataattcac tgtaatgatg taattcttga 1620
 ataataaatc tgacaaaata tt 1642

<210> 76
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>

<308> NM_001548

<400> 76
 gtatcaattc tttgaagaaa ttggttttta ggaaacttcg gagaaaggca ttagatctgg 60

<210> 77
 <211> 3344
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001605

<400> 77

ggtacagctg	cgcgctctg	ggaataggtg	cagcgggccc	ttggcggggg	actctgaggg	60
aggagctggg	gacggcgacc	ctaggagagt	tctttggggg	gactttcaag	atggactcta	120
ctctaacagc	aagtgaatc	cggcagcgat	ttatagattt	cttcaagagg	aacgagcata	180
cgtatgttca	ctcgtctgoc	accatcccac	tggatgaccc	cactttgctc	tttgccaatg	240
caggcatgaa	ccagtttaaa	cccattttcc	tgaacacaat	tgacccatct	caccccatgg	300
caaagctgag	cagagctgoc	aatacccaga	agtgcacccg	ggctgggggc	aaacaaaatg	360
acctggacga	tgtgggcaag	gatgtctatc	atcacacctt	cttcgagatg	ctgggctctt	420
ggtcttttgg	agattacttt	aaggaattgg	catgtaagat	ggctctggaa	ctcctcacc	480
aagagtttgg	cattcccatt	gaaagacttt	atgttactta	ctttggcggg	gatgaagcag	540
ctggcttaga	agcagatctg	gaatgcaaac	agatctggca	aaatttgggg	ctggatgaca	600
ccaaaatcct	cccaggcaac	atgaaggata	acttctggga	gatgggtgac	acgggcccct	660
gtggctcctg	cagttagatc	cactacgacc	ggatttgggtg	tcgggacgcc	gcacatcttg	720
tcaaccagga	cgacccta	gtgctggaga	tctggaacct	tgtgttcac	cagtataaca	780
gggaagctga	tgtcattctg	aaacctcttc	ccaagaaaa	cattgacaca	gggatgggcc	840
tggaaacgact	ggtatctgtg	ctgcagaata	agatgtccaa	ctatgacact	gacctttttg	900
tcccttactt	tgaagccatt	cagaagggca	caggtgcccg	accatacact	gggaaagtgtg	960
gtgctgagga	tgccgatggg	attgacatgg	cctaccgggt	gctggctgac	catgctcgga	1020
ccatcactgt	ggcactgggt	gatggtggcc	ggcctgacaa	cacagggcgt	ggatatgtgt	1080
tgagacggat	tctccgccga	gctgtccgat	acgcccata	aaagctcaat	gccagcaggg	1140
gcttctttgc	tacgttagtg	gatgttgtcg	tccagtcctt	gggagatgca	tttcctgagc	1200
tgaagaagga	cccagacatg	gtgaaggaca	tcattaatga	agaagaggtg	cagtttctca	1260
agactctcag	cagagggcgt	cgcacccctg	acaggaata	tcagagcctg	ggagacagca	1320
agaccattcc	cggagacact	gcttggctcc	tctatgacac	ctatgggttt	ccagtggatc	1380
tgactggact	gattgctgaa	gagaagggcc	tgggtgtaga	catggatggc	tttgaagagg	1440
agaggaaact	ggcccagctg	aaatcacagg	gcaaggagc	tgggtgggaa	gacctcatta	1500
tgtctggacat	ttacgctatc	gaagagctcc	gggcacgggg	tctggaggtc	acagatgatt	1560
ccccaaagta	caattaccat	ttggactcca	gtggttagcta	tgtattttgag	aacacagtgg	1620
ctacgggtgat	ggctctgcgc	agggagaaga	tgttcgtgga	agaggtgtcc	acaggccagg	1680
agtgtggagt	ggctctggac	aagacctgtt	tctatgctga	gcaaggaggc	cagatctatg	1740
acgaaggcta	cctggtgaag	gtggatgaca	gcagtgaaga	taaaacagag	tttacagtga	1800
agaatgctca	ggtccgagga	gggtatgtgc	tacacattgg	aaccatctac	ggtgacctga	1860
aagtggggga	tcaggtctgg	ctgtttattg	atgagccccg	acgaagaccc	atcatgagca	1920
accaacacagc	tacgcacatt	ctgaacttcg	ccctgcgctc	agtgtctggg	gaagctgacc	1980
agaaaggctc	attggttgct	cctgaccgcc	tcagatttga	ctttactgcc	aagggagcca	2040
tgtccaccca	acagatcaag	aaggctgaag	agattgctaa	tgagatgatt	gaggcagcca	2100
aggccgtcta	tacccaggat	tgccccctgg	cagcagcgaa	agccatccag	ggcctacggg	2160
ctgtgtttga	tgagacctat	cctgaccctg	tgcgagtcgt	ctccattggg	gtcccggtgt	2220
ccgagttgct	ggatgacccc	tctgggcctg	ctggctccct	gacttctgtt	gagttctgtg	2280
ggggaacgca	cctgcggaac	tcgagtcatg	caggagcttt	tgtgatcgtg	acggaagaag	2340
ccattgccaa	gggtatccgg	aggattgtgg	ctgtcacagg	tgccgaggcc	cagaaggccc	2400
tcaggaaagc	agagagcttg	aagaaatgtc	tctctgtcat	ggaagccaaa	gtgaaggctc	2460
agactgctcc	aaacaaggat	gtgcagaggg	agatcgctga	ccttggagag	gccctggcca	2520
ctgcaagtcat	cccccaagg	cagaaggatg	aattgcggga	gactctcaaa	tccctaaaga	2580
aggatcatgga	tgacttggac	cgagccagca	aagccgatgt	ccagaaacga	gtgttagaga	2640
agacgaagca	gttcatcgac	agcaacccca	accagcctct	tgtcatcctg	gagatggaga	2700
gcggcgccctc	agccaaaggcc	ctgaatgaag	ccttgaagct	cttcaagatg	cactcccctc	2760
agacttctgc	catgctcttc	acgggtggaca	atgaggctgg	caagatcacg	tgccctgtgtc	2820
aagtcctccca	gaatgcagcc	aatcggggct	taaaagccag	cgagtgggtg	cagcaggtgt	2880
caggcttgat	ggacggtaaa	gggtgtgtgc	atgagtgtgc	tgcacaggcc	acaggcaaga	2940
acgttggctg	cctgcaggag	gcgtgcagc	tgccacttc	cttcgccag	ctgcgcctcg	3000
gggatgtaaa	gaactgagtg	gggaaggagg	aggctccac	tggatccatc	cgtccagcca	3060

```

agagctcttc atctgctaca agaacatttg aatcttgga cctttaaaga gcccctccta 3120
acctcagcagt aactggaaca cacttgggag cagtcctatg tctcagtgcc ccttaaattt 3180
ctgccctgag cctccacgt cagtgccatc ggtctagaac cactaaccoc gcattgctgt 3240
tgatcgtcac gctcgcatct atagataacg gctctccaga cctgagcttt ccgcgtcagc 3300
aagtaggaat cgtttttgct gcagagaata aaaggaccac gtgc 3344

```

<210> 78
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001605

<400> 78
 gccaaagagct cttcatctgc tacaagaaca tttgaatctt gggaccttta aagagcccct 60

<210> 79
 <211> 417
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001645

<400> 79
 acctcccaac caagccctcc agcaaggatt caggagtgcc cctcgggcct cgccatgagg 60
 ctcttctctgt cgctcccggg cctggtggtg gttctgtcga tcgtcttga aggccagcc 120
 ccagcccagg ggaccccaga cgtctccagt gccttgata agctgaagga gtttgaaac 180
 aacttgagg acaaggctcg ggaactcatc agccgcatca aacagagtga actttctgcc 240
 aagatgcggg agtggttttc agagacattt cagaaagtga aggagaaact caagattgac 300
 tcatgaggac ctgaagggtg acatccagga ggggcctctg aaatttcca caccagcg 360
 cctgtgctga ggactccgc catgtggccc caggtgccac caataaaaat cctaccg 417

<210> 80
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001645

<400> 80
 aaacagagtg aactttctgc caagatgcgg gagtggtttt cagagacatt tcagaaagtg 60

<210> 81
 <211> 1389
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_001809

<400> 81
 cgcggaactt tgccaagcac cggctcatgt gaggctcgcg gcacagcgtt ctctgggctc 60
 cccagaagcc agcctttcgc tcccggaccc ggcagccga gcaggagccg tgggaccggg 120
 cgccagcacc ctctgcggcg tgtcatgggc ccgcgcgcgc ggagccgaaa gcccgaggcc 180
 ccgaggaggc gcagcccag cccgaccccg acccccgcc cctcccgccg gggccctcc 240
 ttaggcgctt cctcccatca acacagtcgg cggagacaag gttggctaaa ggagatccga 300
 aagcttcaga agagcacaca cctcttgata aggaagctgc ccttcagccg cctggcaaga 360

```

gaaatatgtg ttaaattcac tcgtggtgtg gacttcaatt ggcaagccca ggccctattg 420
gccctacaag aggcagcaga agcatttcta gttcatctct ttgaggacgc ctatctcctc 480
accttaccatg caggccgagt tactctcttc ccaaaggatg tgcaactggc ccggaggatc 540
cggggccttg aggagggact cggtgagct cctgcaccca gtgtttctgt cagtctttcc 600
tgctcagcca ggggggatga taccggggac tctccagagc catgactaga tccaatggat 660
tctgcgatgc tgtctggact ttgctgtctc tgaacagtat gtgtgtgttg ctttaaataa 720
ttttcttttt tttgagaagg agaagactgc atgactttcc tctgtaacag aggtaataa 780
tgagacaatc aacaccgttc caaaggcctg aaaataatth tcaataaaag agactccaag 840
gttgacttta gtttgtgagt tactcatgtg actatttgag gatthttgaaa acatcagatt 900
tgctgtggta tgggagaaaa ggttatgtac ttattattht agctctttct gtaataatth 960
cattttttac catatgtaca tttgtacttt tattttacac ataagggaag aaataagacc 1020
actttgagca gttgcctgga aggtctggga tttccatcat atagacctct gcccttcaga 1080
gtagcctcac cattagtggc agcatcatgt aactgagtgg actgtgcttg tcaacggatg 1140
tgtagcttht cagaaactta attggggatg aatagaaaac ctgtaagctt tgatgttctg 1200
gttacttcta gtaaatctct gtcaaaatca attcagaaat tctaacttgg agaatttaac 1260
atthtactct tgtaaatcat agaagatgta tcataacagt tcagaattht aaagtacatt 1320
ttcgatgctt ttatgggtat ttttgtagtt tctttgtaga gagataataa aaatcaaaat 1380
atttaatga 1389

```

<210> 82

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001809

<400> 82

```

ggggatgaat agaaaacctg taagctttga tgttctgggt acttctagta aattcctgtc 60

```

<210> 83

<211> 2205

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001909

<400> 83

```

gcgcacgccc gccgcgcccc cgtgaccggt ccgggtgcaa acacgcgggt cagctgatcc 60
ggcccaactg cggcgctatc ccggctataa gcgcacggcc tcggcgaccc tctccgaccc 120
ggccgcgccg gccatgcagc cctccagcct tctgcgctc gccctctgcc tgctggctgc 180
accgcctcc gcgctcgtca ggatcccgtc gcacaagttc acgtccatcc gccggaccat 240
gtcggagggt gggggctctg tggaggacct gattgccaaa ggccccgtct caaagtactc 300
ccaggcggtg ccagccgtga ccgaggggcc cattcccag gtgctcaaga actacatgga 360
cgcccagtag tacggggaga ttggcatcgg gacgcccccc cagtgttcca cagtcgtctt 420
cgacacgggc tctccaacc tgtgggtccc ctccatccac tgcaaaactg tggacatcgc 480
ttgctggatc caccacaagt acaacagcga caagtcagc acctacgtga agaattgtac 540
ctcgtttgac atccactatg gctcgggcag cctctccggg tacctgagcc aggacactgt 600
gtcggtgccc tgccagtcag cgtcgtcagc ctctgccctg ggcggtgtca aagtggagag 660
gcaggtcttt ggggaggcca ccaagcagcc aggcattacc ttcatcgag ccaagttcga 720
ttgcatctgt ggcattggct acccccgcct ctccgtcaac aacgtgctgc ccgtcttoga 780
caacctgatg cagcagaagc tgggtggacca gaacatcttc tcttctacc tgagcaggga 840
cccagatgag cagcctgggg gtgagctgat gctgggtggc acagactcca agtattacaa 900
gggttctctg tctacctga atgtcacccg caaggcctac tggcaggctc acctggacca 960
ggtggagggt gccagcgggc tgaccctgtg caaggagggc tgtgaggcca ttgtggacac 1020
aggcacttcc ctcatggtgg gcccggtgga tgaggtgcgc gagctgcaga aggccatcgg 1080
ggccgtgccg ctgattcagg gcgagtagat gatccccgtg gagaagggtg ccaccctgcc 1140
cgcgatcaca ctgaagctgg gaggcaaagg ctcaaagctg tccccagagg actacacgct 1200
caagggtgag caggccggga agaccctctg cctgagcggc ttcatgggca tggacatccc 1260

```

```

gccacccagc gggccactct ggatcctggg cgacgtcttc atcggcgcgt actacactgt 1320
gtttgaccgt gacaacaaca ggggtgggctt cgccgaggct gccgcctct agttcccaag 1380
gcgtccgcgc gccagcacag aaacagagga ggtcccaga gcaggaggcc cctggcccag 1440
cggccctcc cacacacacc cacacactcg cccgccact gtccctggcg ccctggaagc 1500
cggcgcccca agcccgaatt gctgttttgt tctgtggtt tccctccct gggttcagaa 1560
atgtgcctg cctgtctgtc tctccatctg tttggtggg gtagagctga tccagagcac 1620
agatctgttt cgtgcattgg aagacccac ccaagcttg cagccgagct cgtgtatcct 1680
ggggtccct tcatctccag ggagtcacct ccccgccct accagcgcc gctgggctga 1740
gccctaccc cacaccaggc cgtcctcccg ggccctccct tggaaacctg ccctgcctga 1800
gggcccctct gccagcttg ggcccagctg ggctctgcca ccctacctgt tcagtgtccc 1860
gggcccgttg aggatgaggc cgctagaggc ctgaggatga gctggaagga gtgagagggg 1920
acaaaaccca ccttgttga gctgcaggg tgggtgctgg actgagccag tcccaggggc 1980
atgtattggc ctggaggtgg ggttgggatt gggggctggt gccagccttc ctctgcagct 2040
gacctctgtt gtcctccct tgggcggctg agagccccag ctgacatgga aatacagttg 2100
ttggcctccg gcctccctc aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2160
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa 2205

```

```

<210> 84
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_001909

```

```

<400> 84
tctgtttggt gggggtagag ctgatccaga gcacagatct gtttcgtgca ttggaagacc 60

```

```

<210> 85
<211> 817
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_002038

```

```

<400> 85
gaaccgttta ctgcgtgctg tgcccatcta tcagcaggct ccgggctgaa gattgcttct 60
cttctctoct ccaaggtcta gtgacggagc ccgcgcgcgg cgccaccatg cggcagaagg 120
cggatctcgt tttcttgtgc tacctgctgc tcttcacttg cagtggggtg gaggcaggta 180
agaaaaagtg ctcgagagc tcggacagcg gctccgggtt ctggaaggcc ctgaccttca 240
tggccgtcgg aggaggactc gcagtcgccg ggctgcccg cgtgggcttc accggcgccg 300
gcctcgcgcc caactcgggt gctgcctcgc tgatgagctg gtctgcgata ctgaatgggg 360
gcggcgtgcc cgccgggggg ctagtggcca cgctgcagag cctcggggct ggtggcagca 420
gcgtcgtcat aggtaatatt ggtgcctga tgggctacgc caccacaaag tatctcgata 480
gtgaggagga tgaggagtag ccagcagctc ccagaacctc ttcttcttc ttggcctaac 540
tcttcagtt aggatctaga actttgcctt tttttttttt tttttttttt tttgagatgg 600
gttctcacta tattgtccag gctagagtgc agtggctatt cacagatgcg aacatagtac 660
actgcagcct ccaactccta gcctcaagt atcctcctgt ctcaacctcc caagtaggat 720
tacaagcatg cgccgacgat gccagaatc cagaactttg tctatcactc tcccaacaa 780
cctagatgtg aaaacagaat aaacttcacc cagaaaa 817

```

```

<210> 86
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_002038

```

<400> 86

agctcccaga acctcttctt ccttcttggc ctaactcttc cagttaggat ctagaacttt 60

<210> 87

<211> 1283

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002046

<400> 87

```

ctctctgctc ctctgttctg acagtcagcc gcatcttctt ttgcgtcgcc agccgagcca 60
catcgctcag acaccatggg gaaggtgaag gtccgagtcac acggatttgg tcgtattggg 120
cgctgtgtca ccagggtctg ttttaactct ggtaaagtgg atattgttgc catcaatgac 180
cccttcattg acctcaacta catggtttac atgttccaat atgattccac ccatggcaaa 240
ttccatggca ccgtcaaggc tgagaacggg aagcttgtca tcaatggaaa tcccatcacc 300
atcttcagg agcgagatcc ctccaaaatc aagtggggcg atgctggcgc tgagtacgtc 360
gtggagtcca ctggcgtctt caccaccatg gagaaggctg gggctcattt gcagggggga 420
gccaaaaggg tcatcatctc tgccccctct gctgatgccc ccatgttcgt catgggtgtg 480
aaccatgaga agtatgacaa cagcctcaag atcatcagca atgcctcctg caccaccaac 540
tgcttagcac ccctggccaa ggatcatccat gacaactttg gtatcgtgga aggactcatg 600
accacagtcc atgccatcac tgccaccag aagactgtgg atggcccctc cgggaaactg 660
tggcgtgatg gccgcggggc tctccagaac atcatcctg cctctactgg cgctgccaa 720
gctgtggggc aggtcatccc tgagctgaac ggggaagctc ctggcatggc cttcgtgtc 780
cccactgcca acgtgtcagt ggtggacctg acctgccgtc tagaaaaacc tgccaaatat 840
gatgacatca agaaggtggt gaagcaggcg tcggaggggc ccctcaaggg catcctgggc 900
tacactgagc accaggtggt ctctctgtac ttcaacagcg acaccactc ctccaccttt 960
gacgtctggg ctggcattgc cctcaacgac cactttgtca agctcatttc ctggtatgac 1020
aacgaatttg gctacagcaa cagggtggtg gacctcatg cccacatggc ctccaaggag 1080
taagaccctt ggaccaccag cccagcaag agcacaagag gaagagagag accctcactg 1140
ctggggagtc cctggccacac tcagtcccc accacactga atctcccctc ctacagttg 1200
ccatgtagac cccttgaaga ggggaggggc ctaggggacc gcaccttgtc atgtaccatc 1260
aataaagtac cctgtgtcacc acc 1283

```

<210> 88

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002046

<400> 88

ctcaacgacc actttgtcaa gctcatttcc tggtatgaca acgaatttgg ctacagcaac 60

<210> 89

<211> 1610

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002061

<400> 89

```

ggcacgaggg tcgggcccga gtagccggag ccggagccgc agccaccggg gccttctctt 60
cccgcgcggc ccagccggcc gtccggcctc cctcggggcc gagcgagac caggctccag 120
ccgcgcggcg ccggcagcct cgcgtccctc ctccgggtctc tctcgggctt cgggcaccgc 180
gtcctgtggg cggccgcctg cctgcccgcg cggccgcagc cccttgctg cgggcccctg 240
ggcggcccgt gccatgggca ccgacagcgc cgcggccaag gcgctcctgg cgcgggcccg 300
caccctgcac ctgcagacgg ggaacctgct gaactggggc cgcctgcgga agaagtggcc 360

```

```

gtccacgcac agcgaggagc ttcattgattg tatccaaaaa accttgaatg aatggagttc 420
ccaaatcaac ccagatttgg tcagggagtt tccagatgtc ttggaatgca ctgtatctca 480
tgcagtagaa aagataaatc ctgatgaaag agaagaaatg aaagtctctg caaaactgtt 540
cattgtagaa tcaaactctt catcatcaac tagaagtgc gttgacatgg cctgttcagt 600

ccttggagtt gcacagctgg attctgtgat cattgcttca cctcctattg aagatggagt 660
taatctttcc ttggagcatt tacagcctta ctgggaggaa ttagaaaact tagttcagag 720
caaaaagatt gttgccatag gtacctctga tctagacaaa acacagtttg aacagctgta 780
tcagtgggca caggtaaaac caaatagtaa ccaagttaat cttgcctcct gctgtgtgat 840
gccaccagat ttgactgcat ttgctaaaca atttgacata cagctgttga ctcacaatga 900
tccaaaagaa ctgctttctg aagcaagttt ccaagaagct cttcaggaaa gcattcctga 960
cattcaagcg cacgagtggg tgccgctgtg gctactgcgg tattcgggtca ttgtgaaaag 1020
tagaggaatt atcaaatcaa aaggctacat ttacaagct aaaagaaggg gttcttaact 1080
gacttaggag cataacttac ctgtaatttc cttcaatatg agagaaaatt gagatgtgta 1140
aaatctagtt actgcctgta aatgggtgta ttgaggcaga tattctttcg tcatatttga 1200
cagtatgttg tctgtcaagt tttaaatact tatcttgctt ccatatcaat ccattctcat 1260
gaacctctgt attgctttcc ttaaactatt gttttctaata tgaaattgtc tataaagaaa 1320
atacttgcaa tatatttttc ctttattttt atgactaata taaatcaaga aaatttggtg 1380
ttagatata tttggcctag gtatcagggt aatgtatata catatttttt atttccaaaa 1440
aaaattcatt aattgcttct taactcttat tataaccaag caatttaatt acaattgtta 1500
aaactgaaat actggaagaa gatatttttc ctgtcattga tgagatata cagagtaact 1560
ggagtagctg ggatttacta gtagtgtaaa taaaattcac tcttcaatac 1610

```

<210> 90

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002061

<400> 90

```

ctgacttagg agcataactt acctgtaatt tccttcaata tgagagaaaa ttgagatgtg 60

```

<210> 91

<211> 873

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002106

<400> 91

```

cgcagtttga atcgcggtgc gacgaaggag taggtggtgg gatctcaccg tgggtccgat 60
tagccttttc tctgccttgc ttgcttgagc ttcagcggaa ttcgaaatgg ctggcggtta 120
ggctggaaaag gactccggaa aggccaagac aaaggcgggt tcccgctcgc agagagccgg 180
cttgacagttc ccagtgggcc gtattcatcg acacctaaaa tctaggacga ccagtcatgg 240
acgtgtgggc gcgactgccg ctgtgtacag cgcagccatc ctggagtacc tcaccgcaga 300
ggtacttgaa ctggcaggaa atgcatcaaa agacttaaag gtaaagcgta ttaccctctg 360
tcaacttgcaa cttgctatct gtggagatga agaattggat tctctcatca aggctacaat 420
tgctggtggt ggtgtcatct cacacatcca caaatctctg attgggaaga aaggacaaca 480
gaagactgtc taaaggatgc ctggattcct tgttatctca ggactctaaa tactctaaca 540
gctgtccagt gttggtgatt ccagtggact gtatctctgt gaaaaacaca atttgcctt 600
tttgtaattc tatttgagca agttggaagt ttaattagct ttccaacca ccaaatttct 660
gcattcgagt cttaaccata tttaagtgtt actgtggctt caaagaagct attgattctg 720
aagtagtggg ttttgattga gttgactgtt tttaaaaaac tgtttggtt ttaattgtga 780
tgcagaagtt atagtaacaa acatttggtt ttgtacagac attatttcca ctctggtgga 840
taagttcaat aaaggtcata tcccaaacta aaa 873

```

<210> 92

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002106

<400> 92

cgagtcttaa ccatatttaa gtgttactgt ggcttcaaag aagctattga ttctgaagta 60

<210> 93

<211> 4204

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002205

<400> 93

caggacaggg	aagagcgggc	gctatgggga	gccggacgcc	agagtcccct	ctccacgccg	60
tgcagctgcg	ctggggcccc	cgccgcgcac	ccccgctcgt	gccgctgctg	ttgctgctcg	120
tgccgcgcc	acccagggtc	gggggcttca	acttagacgc	ggaggcccca	gcagtactct	180
cggggcccc	gggctccttc	ttcggattct	cagtggagtt	ttaccggccg	ggaacagacg	240
gggtcagtg	gctggtggga	gcacccaagg	ctaataccag	ccagccagga	gtgctgcagg	300
gtggtgctgt	ctacctctgt	ccttgggggtg	ccagccccac	acagtgcacc	cccattgaat	360
ttgacagcaa	aggctctcgg	ctcctggagt	cctcaactgtc	cagctcagag	ggagaggagc	420
ctgtggagta	caagtccttg	cagtggttcg	gggcaacagt	tcgagcccat	ggctcctcca	480
tcttgcatg	cgctccactg	tacagctggc	gcacagagaa	ggagccactg	agcgaccccc	540
tgggcacctg	ctacctctcc	acagataact	tcacccgaat	tctggagtat	gcaccctgcc	600
gctcagattt	cagctgggca	gcaggacagg	gttactgcc	aggaggcttc	agtcccgagt	660
tcaccaagac	tggccgtgtg	gttttaggtg	gaccaggaag	ctatttcttg	caaggccaga	720
tcctgtctgc	cactcaggag	cagattgcag	aatcttatta	ccccaggtac	ctgatcaacc	780
tggttcaggg	gcagctgcag	actcgccagg	ccagttccat	ctatgatgac	agctacctag	840
gatactctgt	ggctgttggt	gaattcagtg	gtgatgacac	agaagacttt	gttgotggtg	900
tgcccaaagg	gaacctcact	tacggctatg	tcaccatcct	taatggctca	gacattcgat	960
ccctctacaa	cttctcaggg	gaacagatgg	cctcctactt	tggtatgca	gtggccgccca	1020
cagacgtcaa	tggggacggg	ctggatgact	tgctgggtggg	ggcaccctctg	ctcatggatc	1080
ggaccctga	cgggcggcct	caggaggtgg	gcagggtcta	cgtctacctg	cagcaccocag	1140
ccggcataga	gcccacgccc	acccttacct	tcaetggcca	tgatgagttt	ggccgatttg	1200
gcagctcctt	gacccccctg	ggggacctgg	accaggatgg	ctacaatgat	gtggccatcg	1260
gggtccctt	tgggtggggag	accagcagg	gagtagtggt	tgtatttcct	gggggcccag	1320
gagggctggg	ctctaagcct	tcccaggttc	tgcagccctt	gtgggcagcc	agccacaccc	1380
cagacttctt	tggctctgcc	cttcgaggag	gccgagacct	ggatggcaat	ggatatcctg	1440
atctgattgt	ggggctcctt	ggtgtggaca	aggctgtggt	atacaggggc	cgccccatcg	1500
tgtccgctag	tgcctccctc	accatcttcc	ccgccatgtt	caaccocagag	gagcggagct	1560
gcagcttaga	ggggaaccct	gtggcctgca	tcaaccttag	cttctgcctc	aatgcttctg	1620
gaaaacacgt	tgttgactcc	attggtttca	cagtggaact	tcagctggac	tggcagaagc	1680
agaagggagg	ggtacggcgg	gcactgttcc	tggcctccag	gcaggcaacc	ctgaccaga	1740
ccctgctcat	ccagaatggg	gctcgagagg	attgcagaga	gatgaagatc	tacctcagga	1800
acgagtca	atttcgagac	aaactctcgc	cgattcacat	cgctctcaac	ttctccttgg	1860
acccccaagc	cccagtggac	agccacggcc	tcaggccagc	cctacattat	cagagcaaga	1920
gccgataga	ggacaaggct	cagatcttgc	tggactgtgg	agaagacaac	atctgtgtgc	1980
ctgacctgca	gctggaagtg	tttggggagc	agaaccatgt	gtacctgggt	gacaagaatg	2040
ccctgaacct	cactttccat	gccagaatg	tgggtgaggg	tggcgccctat	gaggetgagc	2100
ttcgggtcac	cgccccctca	gaggctgagt	actcaggact	cgtcagacac	ccagggaact	2160
tctccagcct	gagctgtgac	tactttgccg	tgaaccagag	ccgcctgctg	gtgtgtgacc	2220
tgggcaaccc	catgaaggca	ggagccagtc	tgtgggggtg	ccttcggttt	acagtccctc	2280
atctccggga	cactaagaaa	accatccagt	ttgacttcca	gatcctcagc	aagaatctca	2340
acaactcgca	aagcgacgtg	gtttccttcc	ggctctccgt	ggaggctcag	gcccagggtca	2400
ccctgaacgg	tgtctccaag	cctgaggcag	tgctattccc	agtaagcgac	tggcatcccc	2460
gagaccagcc	tcagaaggag	gaggacctgg	gacctgctgt	ccaccatgtc	tatgagctca	2520

tcaaccaagg	ccccagctcc	attagccagg	gtgtgctgga	actcagctgt	ccccaggctc	2580
tggaaggtca	gcagctccta	tatgtgacca	gagttacggg	actcaactgc	accaccaatc	2640
accccatata	cccaaagggc	ctggagttgg	atcccagggg	ttccctgcac	caccagcaaa	2700
aacgggaagc	tccaagccgc	agctctgctt	cctcgggacc	tcagatcctg	aaatgcccg	2760
aggctgagtg	tttcaggctg	cgctgtgagc	tggggcccct	gcaccaacaa	gagagccaaa	2820
gtctgcagtt	gcatttccga	gtctgggcca	agactttctt	gcagcgggag	caccagccat	2880
ttagcctgca	gtgtgaggct	gtgtacaaa	ccctgaagat	gccctaccga	atcctgcctc	2940
ggcagctgcc	ccaaaaagag	cgtcagggtg	ccacagctgt	gcaatggacc	aaggcagaag	3000
gcagctatgg	cgtcccactg	tggatcatca	tcttagccat	cctgtttggc	ctcctgctcc	3060
taggtctact	catctacatc	ctctacaagc	ttggattctt	caaacgctcc	ctcccatatg	3120
gcaccgccat	ggaaaaagct	cagctcaaac	ctccagccac	ctctgatgcc	tgagtctctc	3180
caatttcaga	ctcccattcc	tgaagaacca	gtccccccac	cctcattcta	ctgaaaagga	3240
gggtcttggg	tacttcttga	aggtgctgac	ggccaggagg	aagctcctct	ccccagccca	3300
gagacatact	tgaagggccca	gagccagggg	ggtgaggagc	tggggatccc	tccccccat	3360
gcactgtgaa	ggacccttgt	ttacacatac	cctcttcatg	gatgggggaa	ctcagatcca	3420
gggacagagg	cccagcctcc	ctgaagcctt	tgcatttttg	agagtttcct	gaaacaactg	3480
gaaagataac	taggaaatcc	attcacagtt	ctttgggcca	gacatgccac	aaggacttcc	3540
tgtccagctc	caacctgcaa	agatctgtcc	tcagccttgc	cagagatcca	aaagaagccc	3600
ccagtaagaa	cctggaactt	ggggagttaa	gacctggcag	ctctggacag	ccccaccctg	3660
gtgggccaac	aaagaacact	aactatgcat	ggtgccccag	gaccagctca	ggacagatgc	3720
cacaaggata	gatgctggcc	cagggccaga	gccagctccc	aaggggaatc	agaactcaaa	3780
tggggccaga	tccagcctgg	ggtctggagt	tgatctggaa	cccagactca	gacattggca	3840
ccaatccagg	cagatccagg	actatatttg	ggcctgctcc	agacctgac	ctggaggccc	3900
agttcaccc	gatttaggag	aagccaggaa	tttcccagga	cctgaagggg	ccatgatggc	3960
aacagatctg	gaacctcagc	ctggccagac	acaggccctc	cctgttcccc	agagaaaggg	4020
gagccactg	tcttgggctc	gcagaatttg	ggttctgcct	gccagctgca	ctgatgctgc	4080
ccctcatctc	tctgcccac	ccttcctcca	ccttggcacc	agacaccag	gacttattta	4140
aactctgttg	caagtgcaat	aaatctgacc	cagtgcctcc	actgaccaga	actagaaaaa	4200
aaaa	4204					

<210> 94

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002205

<400> 94

ttggcaccag acaccagga cttattttaa ctctgttgca agtgcaataa atctgaccca 60

<210> 95

<211> 1976

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002266

<400> 95

gccacacggt	ctttgagctg	agtcgaggtg	gaccctttga	acgcagtcgc	cctacagccg	60
ctgattcccc	cgcctcgc	tcccgaggaa	gcccaggccc	gcttcgcagc	tttctccctt	120
tgtctcataa	ccatgtccac	caacgagaat	gctaatacac	cagctgcccg	tcttcacaga	180
ttcaagaaca	agggaaaaga	cagtcacagaa	atgaggcgctc	gcagaataga	ggtcaatgtg	240
gagctgagga	aagctaagaa	ggatgaccag	atgctgaaga	ggagaaatgt	aagctcattt	300
cctgatgatg	ctacttctcc	gctgcaggaa	aaccgcaaca	accagggcac	tgtaaatttg	360
tctgttgatg	acattgtcaa	aggcataaat	agcagcaatg	tggaaaatca	gctccaagct	420
actcaagctg	ccaggaaaact	actttccaga	gaaaaacagc	cccccataga	caacataatc	480
cgggctgggt	tgattccgaa	atttgtgtcc	ttcttgggca	gaactgattg	tagtccattt	540
cagtttgatg	ctgcttgggc	actcactaac	attgcttctg	ggacatcaga	acaaaccaag	600
gctgtggtag	atggaggtgc	catoccagca	ttcatttctc	tgttggcatc	tccccatgct	660

```

cacatcagtg aacaagctgt ctgggctcta ggaacattg cagggtgatgg ctcagtgttc 720
cgagacttgg ttattaaagta cgggtgcagtt gaccactgt tggctctcct tgcagttcct 780
gatatgtcat ctttagcatg tggctactta cgtaatctta cctggacact ttctaactctt 840
tgccgcaaca agaatcctgc acccccgata gatgctgttg agcagattct tctactctta 900
gttcggctcc tgcacatga tgatccagaa gtgttagcag atacctgctg ggctatttcc 960
taccttactg atggtccaaa tgaacgaatt ggcattggtg tgaaaacagg agttgtgccc 1020
caacttgtga agcttctagg agcttctgaa ttgccaattg tgactcctgc cctaagagcc 1080
atagggaata ttgtcactgg tacagatgaa cagactcagg ttgtgattga tgcaggagca 1140
ctcgccgtct ttcccagcct gctcaccaac cccaaaacta acattcagaa ggaagctacg 1200
tggacaatgt caaacatcac agccggccgc caggaccaga tacagcaagt tgtgaatcat 1260
ggattagtcc cattccttgt cagtgttctc tctaaggcag attttaagac acaaaaggaa 1320
gctgtgtggg ccgtgaccaa ctataccagt ggtggaacag ttgaacagat tgtgtacctt 1380
gttactgtg gcataataga accgttgatg aacctcttaa ctgcaaaaga taccaagatt 1440
attctgggta tccctgatgc catttcaa atctttcagg ctgctgagaa actaggtgaa 1500
actgagaaac ttagtataat gattgaagaa tgtggaggct tagacaaaat tgaagctcta 1560
caaaaccatg aaaaatgagtc tgtgtataag gcttcgttaa gcttaattga gaagtatttc 1620
tctgtagagg aagaggaaga tcaaaacggt gtaccagaaa ctacctctga aggctacact 1680
ttccaagtgc aggatggggc tcctgggacc tttaactttt agatcatgta gctgagacat 1740
aaatttggtg tgtactacgt ttggtatttt gtcttattgt ttctctacta agaactcttt 1800
cttaaatgtg gtttgttact gtagcacttt ttacactgaa actatacttg aacagttcca 1860
actgtacata catactgtat gaagcttgct ctctgactag gtttctaatt tctatgtgga 1920
atttcctatc ttgcagcatc ctgtaaataa acattcaagt ccacccttaa aaaaaa 1976

```

<210> 96

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002266

<400> 96

```

tgagtctgtg tataaggctt cgtaaagctt aattgagaag tatttctctg tagaggaaga 60

```

<210> 97

<211> 1145

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002346

<400> 97

```

gtcccgcca gcccggtcc agagcgcgcg aggttcgggg agctccgcca ggctgctggt 60
acctgctcc gcccgcgag caggacaggc tgctttggtt tgtgacctcc aggaggagc 120
gccatcctct ccagaatgaa gatcttcttg ccagtgcctg tggctgcctt tctgggtgtg 180
gagcgagcca gctcgctgat gtgcttctcc tgcttgaacc agaagagcaa tctgtactgc 240
ctgaagccga ccatctgctc cgaccaggac aactactgcg tgactgtgtc tgctagtgcc 300
ggcattggga atctcgtgac atttggccac agcctgagca agacctgttc cccggcctgc 360
cccatcccag aaggcgtcaa tgttgggtgtg gcttccatgg gcatacagctg ctgccagagc 420
tttctgtgca atttcagtgc ggccgatggc gggtgcgggg caagcgtcac cctgctgggt 480
gccgggctgc tgcctgagct gctgcgggcc ctgctgcggg ttggcccttg accgccaga 540
ccctgtcccc cgatccccc gctcaggaag gaaagcccag ccctttctgg atcccacagt 600
gtatgggagc ccctgactcc tcacgtgcct gatctgtgcc cttggtccca ggtcaggccc 660
acccctgca cctccacctg ccccagcccc tgctctgccc caagtgggcc agctgccttc 720
acttctgggg tggatgatgt gaccttccct gggggactgc ggaaggagc agggttccct 780
ggagtcttac ggtccaacat cagaccaagt cccatggaca tgctgacagg gtcccaggg 840
agaccgtgtc agtagggatg tgtgcctggc tgtgtacgtg ggtgtgcagt gcacgtgaga 900
gcacgtggcg gcttctgggg gccatgtttg gggaggagg tgtgccagca gcctggagag 960
cctcagtcct tgtagcccc tgccctggca cagctgcatg cacttcaagg gcagcctttg 1020
ggggttgggg tttctgccac ttccgggtct aggccctgcc caaatccagc cagtccctgc 1080

```

ccagcccacc cccacattgg agccctcctg ctgctttggt gcctcaaata aatacagatg 1140
tcccc 1145

<210> 98
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002346

<400> 98
ggttccctgg agtcttacgg tccaacatca gaccaagtcc catggacatg ctgacagggt 60

<210> 99
<211> 1390
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002358

<400> 99
gggaagtgct gttggagccg ctgtggttgc tgtccgcgga gtggaagcgc gtgcttttgt 60
tttgtccct ggccatggcg ctgcagctct ccgaggagca gggaatcacc ctgcgcggga 120
gcgcggaat cgtggccgag ttcttctcat tcggcatcaa cagcatttta tatcagcgtg 180
gcataatatcc atctgaaacc ttactcgag tgcagaaata cggactcacc ttgcttgtaa 240
ctactgatct tgagctcata aaatacctaa ataatgtggt ggaacaactg aaagattggt 300
tatacaagtg ttcagttcag aaactggtt tagttatctc aaatattgaa agtggtgagg 360
tcttggaag atggcagttt gatattgagt gtgacaagac tgcaaaagat gacagtgcac 420
ccagagaaaa gtctcagaaa gctatccagg atgaaatccg ttcagtgatc agacagatca 480
cagctacggg gacatttctg ccactgttgg aagtttcttg ttcatttgat ctgctgattt 540
atacagacaa agatttggtt gtacctgaaa aatgggaaga gtcgggacca cagtttatta 600
ccaattctga ggaagtccgc ctctgctcat ttactactac aatccacaaa gtaaatagca 660
tggtggccta caaaattcct gtcaatgact gaggatgaca tgaggaaaat aatgtaattg 720
taattttgaa atgtggtttt cctgaaatca ggatcatctat agttgatatg ttttatttca 780
ttggttaatt ttacatgga gaaaaccaa atgatactta ctgaactgtg tgaattgtt 840
cctttatttt ttggtacct atttgactta ccatggagtt aacatcatga atttattgca 900
cattgttcaa aagggaaccag gaggtttttt tgtcaacatt gtgatgtata ttcctttgaa 960
gatagtaact gtagatggaa aaacttgtgc tataaagcta gatgctttcc taaatcagat 1020
gttttggcca agtagtttga ctacgtatag gtagggagat atttaagtat aaaatacaac 1080
aaaggaaagtc taaatattca gaatctttgt taaggctcctg aaagtaactc ataactata 1140
aacaatgaaa tattgctgta tagctccttt tgacctcat ttcattgtata gttttcccta 1200
ttgaatcagt ttccaattat ttgacttta tttatgtaac ttgaacctat gaagcaatgg 1260
atattgtgac tgtttaatgt tctgtgatac agaactctta aaaatgtttt ttcattgtgtt 1320
ttataaaatc aagttttaag tgaaagtggag gaaataaagt taagtttggt ttaaaaaaaa 1380
aaaaaaaaa 1390

<210> 100
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002358

<400> 100
atgctttcct aaatcagatg ttttgggtcaa gtagtttgac tcagtatagg tagggagata 60

<210> 101
<211> 1821

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002422

<400> 101

```

acaaggaggc aggcaagaca gcaaggcata gagacaacat agagctaagt aaagccagtg 60
gaaatgaaga gtcttccaat cctactgttg ctgtgcgtgg cagtttgctc agcctatcca 120
ttggatggag ctgcaagggg tgaggacacc agcatgaacc ttgttcagaa atatctagaa 180
aactactacg acctcaaaaa agatgtgaaa cagtttggtt ggagaaagga cagtggctct 240
gttggttaaaa aaatccgaga aatgcagaag ttccttggat tggagggtgac ggggaagctg 300
gactccgaca ctctggaggt gatgcgcaag cccaggtgtg gagttcctga tgttggtcac 360
ttcagaacct ttcctggcat cccgaagtgg aggaaaaccc accttacata caggattgtg 420
aattatacac cagatttgcc aaaagatgct gttgattctg ctgttgagaa agctctgaaa 480
gtctgggaag aggtgactcc actcacattc tcaggctgt atgaaggaga ggctgatata 540
atgatctctt ttgcagttag agaaccatgga gacttttacc cttttgatgg acctggaat 600
gttttggccc atgcctatgc ccctgggcca gggattaatg gagatgcca ctttgatgat 660
gatgaacaat ggacaaagga tacaacaggg accaatat tctcgttg tgctcatgaa 720
attggccact ccctgggtct ctttactca gccaacactg aagctttgat gtaccactc 780
tatcactcac tcacagacct gactcggttc cgctgtctc aagatgat aaatggcatt 840
cagtcctct atggacctcc ccctgactcc cctgagaccc ccctgggtacc caggaacct 900
gtccctccag aacctgggac gccagccaac tgtatcctg ctttgcctt tgatgctgtc 960
agcactctga ggggagaaat cctgatctt aaagacaggc acttttggcg caaatccctc 1020
aggaagcttg aacctgaatt gcatttgatc tcttcattt ggccatctc tccttcaggc 1080
gtggatgccg catatgaagt tactagcaag gacctcgtt tcatttttaa aggaaatcaa 1140
ttctgggcca tcagaggaaa tgaggtacga gctggatacc caagaggcat ccacacccta 1200
ggtttccctc caacogtgag gaaaatcgat gcagccatt ctgataagga aaagaacaaa 1260
acataattct ttgtagagga caaatactgg agatttgatg agaagagaaa ttccatggag 1320
ccaggctttc ccaagcaaat agctgaagac tttccaggga ttgactcaaa gattgatgct 1380
gtttttgaag aatttgggtt cttttatttc tttactggat cttcacagtt ggagtttgac 1440
ccaaatgcaa agaaagtgac acacactttg aagagtaaca gctggcttaa ttgttgaaag 1500
agatatgtag aaggcacaat atgggcactt taaatgaagc taataattct tcacctaaag 1560
ctctgtgaat tgaaatgttc gttttctcct gcctgtgctg tgactcgagt cactcaag 1620
ggaacttgag cgtgaatctg tatcttgccg gtcattttta tgttattaca gggcattcaa 1680
atgggctgct gcttagcttg caccttgtca catagagtga tctttcccaa gagaagggga 1740
agcactcgtg tgcaacagac aagtgactgt atctgtgtag actatttgct tatttaataa 1800
agacgatttg tcagttgttt t 1821

```

<210> 102

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002422

<400> 102

```

tgtagaaggc acaatatggg cactttaaat gaagctaata attcttcacc taagtctctg 60

```

<210> 103

<211> 2787

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002462

<400> 103

```

agagcggagg ccgactcca gcaactgcga gggaccgcct tggaccgcag ttgccggcca 60
ggaatcccag tgtcacgggt gacaogcctc cctcgcgcct ttgccgcca cctgctcacc 120

```

```

cagctcaggg gctttggaat tctgtggcca cactgcgagg agatcgggtc tgggtcggag 180
gctacaggaa gactccact ccctgaaatc tggagtgaag aacgcccca tccagccacc 240
attccaagga ggtgcaggag aacagctctg tgataccatt taacttgttg acattacttt 300
tatttgaagg aacgtatat agagcttact ttgcaaagaa ggaagatggt tgtttccgaa 360
gtggacatcg caaaagctga tccagctgct gcatccacc ctctattact gaatggagat 420
gctactgtgg cccagaaaaa tccaggctcg gtggctgaga acaacctgtg cagccagtat 480
gaggagaagg tgcgcccctg catcgacctc attgactccc tgcgggctct aggtgtggag 540
caggacctgg ccctgccagc catcgccgtc atcggggacc agagctcggg caagagctcc 600
gtgttggagg cactgtcagg agttgccctt cccagaggca gcgggatcgt gaccagatgc 660
ccgctggtagc tgaaactgaa gaaacttgtg aacgaagata agtggagagg caaggtcagt 720
taccaggact acgagattga gatttcggat gcttcaggag tagaaaagga aattaataaa 780
gccagaatg ccacgcggg ggaaggaatg ggaatcagtc atgagcta at caccctggag 840
atcagctccc gagatgtccc ggatctgact ctaatagacc ttcttgcat aaccagagt 900
gctgtgggca atcagcctgc tgacattggg tataagatca agacactcat caagaagtac 960
atccagaggc aggagacaat cagcctggtg gtggtcccca gtaatgtgga catcgccacc 1020
acagaggctc tcagcatggc ccaggagggtg gaccccgagg gagacaggac catcggaatc 1080
ttgacgaagc ctgactctgt ggacaaagga actgaagaca aggttgtgga cgtggtgcgg 1140
aaactcgtgt tccacctgaa gaagggttac atgattgtca agtgccggg ccagcaggag 1200
atccaggacc agctgagcct gtccgaagcc ctgcagagag agaagatctt ctttgagaac 1260
caccatatt tcagggatct gctggaggaa ggaaggcca cggttccctg cctggcagaa 1320
aaacttacca gcgagctcat cacacatata tgtaaatctc tgcccctgtt agaaaatcaa 1380
atcaaggaga ctcaccagag aataacagag gagctacaaa agtatggtgt cgacataccg 1440
gaagacgaaa atgaaaaaat gttcttctct atagataaaa ttaatgcctt taatcaggac 1500
atcactgctc ctgatcaagg agaggaaact gtatgggagg aagacattcg gctgtttacc 1560
agactccgac acgagttcca caaatggagt acaataattg aaaacaattt tcaagaaggc 1620
cataaaattt tgagttagaa aatccagaaa tttgaaaatc agtatcgtgg tagagagctg 1680
ccaggctttg tgaattacag gacatttgag acaatcgtga aacagcaaat caaggcactg 1740
gaagagccgg ctgtggatat gctacacacc gtgacggata tgggtccggct tgctttcaca 1800
gatgtttcga taaaaaattt tgaagagttt tttaacctcc acagaaccgc caagtccaaa 1860
attgaagaca ttagagcaga acaagagaga gaaggtgaga agctgatccg cctccacttc 1920
cagatggaac agattgtcta ctgccaggac caggtataca ggggtgcatt gcagaaggtc 1980
agagagaagg agctgtaaga agaaaaaag aagaaatcct gggatttttg ggctttccag 2040
tccagctcgg caacagactc ttccatggag gagatctttc agcacctgat ggccatcac 2100
caggaggcca gcaagcgc atccagccac atccctttga tcatccagtt cttcatgctc 2160
cagacgtacg gccagcagct tcagaaggcc atgctgcagc tcttcagga caaggacacc 2220
tacagctggc tcctgaagga gcggagcgac accagcgaca agcggaagt cctgaaggag 2280
cggcttgca cggctgacga ggctcggcgc cggcttgccc agttcccccg ttaaccacac 2340
tctgtccagc cccgtagacg tgcacgcaca ctgtctgccc cggttcccg gtageccactg 2400
gactgacgac ttgagtgc atagtcaga ctggatagtc cgtctctgct tatccgttag 2460
ccgtggtgat ttagcaggaa gctgtgagag cagtttggtt tctagcatga agacagagcc 2520
ccaccctcag atgcacatga gctggcggga ttgaaggatg ctgtcttcgt actgggaaag 2580
ggattttcag ccctcagaat cgctccacct tgcagctctc cccttctctg tttcctaga 2640
aactgacaca tgctgaacat cacagcttat ttctcattt ttataatgtc cttcacaaa 2700
cccagtggtt taggagcatg agtgccgtgt gtgtgcgtcc tgtcggagcc ctgtctcctc 2760
tctctgtaat aaactcattt ctagcag 2787

```

<210> 104

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002462

<400> 104

actgacacat gctgaacatc acagcttatt tctcatttt tataatgtcc cttcacaaac 60

<210> 105

<211> 2808

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002759

<400> 105

```
gcggcggcgg cggcgcagtt tgctcatact ttgtgacttg cggtcacagt ggcattcagc 60
tccacacttg gtagaaccac aggcacgaca agcatagaaa catcctaaac aatcttcate 120
gaggcatcga ggtccatccc aataaaaaatc aggagaccct ggctatcata gaccttagtc 180
ttcgctggta tactcgctgt ctgtcaacca gcggttgact ttttttaagc cttctttttt 240
ctcttttacc agtttctgga gcaaatcag tttgccttcc tggatttgta aattgtaatg 300
acctcaaaac tttagcagtt cttccatctg actcagggtt gcttctctgg cggctcttcag 360
aatcaacatc cacacttccg tgattatctg cgtgcatttt ggacaaagct tccaaccagg 420
atacgggaag aagaaatggc tggatgatct tcagcagggt tcttcatgga ggaacttaat 480
acataccgtc agaagcaggg agtagtactt aaatatcaag aactgcctaa ttcaggacct 540
ccacatgata ggaggtttac atttcaagtt ataatagatg gaagagaatt tccagaaggt 600
gaaggtagat caaagaagga agcaaaaaat gccgcagcca aattagctgt tgagatactt 660
aataaggaaa agaacggcagt tagtccttta ttattgacaa caacgaattc ttcagaagga 720
ttatccatgg ggaattacat aggccttata aatagaattg cccagaagaa aagactaact 780
gtaaattatg aacagtgtgc atcgggggtg catgggccag aaggatttca ttataaatgc 840
aaaaatgggac agaaagaata tagtattggt acagggttcta ctaaacagga agcaaaacaa 900
ttggccgcta aacttgcata tcttcagata ttatcagaag aaacctcagt gaaatctgac 960
tacctgtcct ctggttcttt tgctactacg tgtgagctcc aaagcaactc ttagtgacc 1020
agcacactcg cttctgaatc atcatctgaa ggtgacttct cagcagatac atcagagata 1080
aattctaaca gtgacagttt aaacagttct tcgttgctta tgaatggtct cagaaataat 1140
caaaggaagg caaaaagatc tttggcacc agatttgacc ttcctgacat gaaagaacaa 1200
aagtatactg tggacaagag gtttggcatg gattttaaag aaatagaatt aattggctca 1260
ggtggatttg gccaaagttt caaagcaaaa cacagaattg acggaagac ttacgttatt 1320
aaacgtgtta aatataataa cgagaaggcg gagcgtgaag taaaagcatt ggcaaaactt 1380
gatcatgtaa atattgttca ctacaatggc tgttgggatg gatttgatta tgatcctgag 1440
accagtgatg attctcttga gagcagtgat tatgatcctg agaacagcaa aaatagttca 1500
aggtaaaaga ctaagtgcct tttcatccaa atggaattct gtgataaagg gaccttgaa 1560
caatggattg aaaaaagaa aggcgagaaa ctagacaaa ttttggtctt ggaactcttt 1620
gaacaaataa caaaaggggt ggattatata cattcaaaaa aattaattca tagagatctt 1680
aagccaagta atatatctt agtagataca aaacaagtaa agattggaga ctttggactt 1740
gtaacatctc tgaanaatga tggaaagcga acaaggagta agggaacttt gcgatacatg 1800
agcccagaac agatttcttc gcaagactat ggaaaggaag tggacctcta cgctttgggg 1860
ctaattcttg ctgaacttct tcatgtatgt gacactgctt ttgaaacatc aaagtttttc 1920
acagacctac gggatggcat catctcagat atatttgata aaaaagaaaa aactcttcta 1980
cagaaattac tctcaagaa acctgaggat cgacctaaac catctgaaat actaaggacc 2040
ttgactgtgt ggaagaaaag cccagagaaa aatgaacgac acacatgtta gagcccttct 2100
gaaaaagtat cctgcttctg atatgcagtt ttccttaaat tatctaaaat ctgctaggga 2160
atatcaatag atattttact tttattttaa tgtttccttt aattttttac tatttttact 2220
aatctttctg cagaaacaga aaggttttct tctttttgct tcaaaaacat tcttacattt 2280
tactttttcc tggctcatct ctttattctt tttttttttt ttaaagacag agtctcgctc 2340
tgttgccag gctggagtgc aatgacacag tcttggctca ctgcaacttc tgctcttgg 2400
gttcaagtga ttctcctgcc tcagcctcct gagtagctgg attacaggca tgtgccacc 2460
acccaactaa tttttgtgtt ttttaataaag acagggtttc accatgttgg ccaggctggt 2520
ctcaaaactc tgacctcaag taatccacct gcctcggcct cccaaagtgc tgggattaca 2580
gggatgagcc accgcgccca gcctcatctc tttgttctaa agatggaaaa accaccccc 2640
aattttcttt ttatactatt aatgaatcaa tcaattcata totattttat aaatttctac 2700
cgcttttagg ccaaaaaaat gtaagatcgt tctctgcctc acatagctta caagccagct 2760
ggagaaatat ggtactcatt aaaaaaaaaa aaaaagtgat gtacaacc 2808
```

<210> 106

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002759

<400> 106
tcgttctctg cctcacatag cttacaagcc agctggagaa atatggtact cattaaaaaa 60

<210> 107
<211> 1678
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002811

<400> 107
aagaaggagg ccgcgcgagg gctgacgaac cggaagaaga ggaactgggc ctgaaagggg 60
accggtgacc gctactgctg ccggtgtttg cgtgtggcag ggagccaggc ctggcgagcg 120
gggtgtgtcg cgatgccgga gctggcagtg cagaagggtg tgggccaccc cctggtgctg 180
ctcagtgtgg tggatcatth caaccgaatc ggcaagggtg gaaaccagaa gcgtgttggt 240
ggtgtgcttt tggggtcatg gcaaaagaaa gtacttgatg tatcgaacag ttttgcagtt 300
ccttttgatg aagatgacaa agacgattct gtatggtttt tagaccatga ttatttggaa 360
aacatgtatg gaatgtttaa gaaagtcaat gccagggaaa gaatagttgg ctggtaccac 420
acaggcccta aactacacaa gaatgacatt gccatcaacg aactcatgaa aagatactgt 480
cctaattccg tattggtcat cattgatgtg aagccgaagg acctagggct gcctacagaa 540
gcgtacattt cagtgaaga agtccatgat gatggaactc caacctcgaa aacatttgaa 600
cacgtgacca gtgaaattgg agcagaggaa gctgaggaag ttggagttga acacttgta 660
cgagatatca aagacacgac ggtgggcact ctgtcccagc ggatcacaaa ccagggtccat 720
ggtttgaagg gactgaactc caagcttctg gatatcagga gctacctgga aaaagtcgcc 780
acaggcaagc tgcccatcaa ccaccagatc atctaccagc tgcaggacgt cttcaacctg 840
ctgccagatg tcagcctgca ggagttcgtc aaggcctttt acctgaagac caatgaccag 900
atggtggtag tgtacttggc ctgcgtgatc cgttccgtgg tcgccctgca caacctcatc 960
aacaacaaga ttgccaaccg ggatgcagag aagaaagaag ggcaggagaa agaagagagc 1020
aaaaaggata ggaaagagga caaggagaaa gataaagata aggaaaagag tgatgtaaa 1080
aaagaggaga aaaaggagaa aaagtaaaac atgtattaaa tagctttttt aatttgtaaa 1140
ttaaaatctt acaaaactaaa tcagtgtgct gctagagggt tctttttcac ttgacatgct 1200
tattagaaag ctgaccacac aagagctctc tgcctccggt cactcttgct gtggtgctac 1260
gtggaagtga atggagactg atctcaaact tgaactgcag ctttcgctgc tgtgagttgg 1320
ggatatgata gtcagctcag gcttcagatt gtatgagaaa aatgaagaga agtcaacaaa 1380
tattttgga ctcttcattc atttatctct aaaaccagga gttgaatttt cctcatcttg 1440
aaagactctt ggggtctggt tctggtatth taaaaaattg ctaagtggaa tgcataaatt 1500
gcattatggt ctctgttaac acgtagagtt cagacccttc tgaactctgt tgataatacc 1560
acaccatggt ctggacccat agctctggca tctcagggg ttgtgatcca gctccatata 1620
ttgtttacct tcaaagatac aattaaatgg cttgattttt aaaaaaaaaa aaaaaaaa 1678

<210> 108
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002811

<400> 108
aaattgctaa gtggaatgca tgaattgcat tatgttctct ggtaacacgt agagttcaga 60

<210> 109
<211> 846
<212> DNA
<213> Homo sapiens

<300>
<308> NM_002888

<400> 109

```

ccacgtccgg ggtgccgagc caactttcct gcgtccatgc agccccgccg gcaacggctg 60
cccgtccctt ggtccggggc cagggggccg cgcgccacgc ccccgctgct cgcgctgctg 120
ctgttgctcg ccccggtggc ggcgcccgcg ggggccgggg gccccgacga ccctgggcag 180
cctcaggatg ctgggggtccc gcgcaggctc ctgcagcaga aggcgcgcgc ggcgcttcac 240
ttcttcaact tccgggtccg ctgcgccagc gcgctgcgag tgctggccga ggtgcaggag 300
ggccgcgcgt ggattaatcc aaaagaggga tgtaaagtgc acgtggtctt cagcacagag 360
cgctacaacc cagagtcttt acttcaggaa ggtgagggac gtttggggaa atgttctgct 420
cgagtgtttt tcaagaatca gaaaccaga ccaaccatca atgtaacttg tacacggctc 480
atcgagaaaa agaaaaagaca acaagaggat tacctgcttt acaagcaa atgaagcaactg 540
aaaaaccctt tggaatatgt cagcatacct gataatcatg gacatatgta tccctctctg 600
agactcatct gggatttggc tttccttgga agctcttacg tgatgtggga aatgacaaca 660
caggtgtcac actactactt ggcacagctc actagtgtga ggcagtgggt aagaaaaacc 720
tgaaaattaa cttgtgccac aagagtta caatcaaagtgg tctccttaga ctgaattcat 780
gtgaacttct aatttcatat caagagttgt aatcacattt atttcaataa atatgtgagt 840
tcctgc 846

```

<210> 110

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002888

<400> 110

```

aaagaaaaga caacaagagg attacctgct ttacaagcaa atgaagcaac tgaaaaaccc 60

```

<210> 111

<211> 1054

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003090

<400> 111

```

gaattccgcg ggaggccacg ggctttccac agcgcggggg aacggggaggc tgcaggatgg 60
tcaagctgac ggcggagctg atcgagcagg cggcgagcag caccaacgcg gtgcgcgacc 120
gggagctgga cctccggggg tataaaattc ccgtcattga aaatctaggt gctacgttag 180
accagtttga tgctattgat ttttctgaca atgagatcag gaaactggat ggttttcctt 240
tgttgagaag actgaaaaca ttgttagtga acaacaacag aatatgccgt ataggtgagg 300
gacttgatca ggctctgccc tgtctgacag aactcattct caccaataat agtctcgtgg 360
aactgggtga tctggaccct ctggcatctc tcaaatcgct gacttaccta agtatocctaa 420
gaaatccggt aaccaataag aagcattaca gattgtatgt gatttataaa gttccgcaag 480
tcagagtact ggatttccag aaagtgaac taaaagagcg tcaggaagca gagaaaatgt 540
tcaagggcaa acggggtgca cagcttgcaa aggatattgc caggagaagc aaaactttta 600
atccaggtgc tggtttgcca actgacaaaa agagaggtgg gccatctcca ggggatgtag 660
aagcaatcaa gaatgccata gcaaatgctt caactctggc tgaagtggag aggctgaagg 720
ggttgctgca gtctggtcag atccctggca gagaacgcag atcagggccc actgatgatg 780
gtgaagaaga gatggaagaa gacacagtca caaacgggtc ctgagcagtg aggcagatgt 840
ataataatga gccctcttgg aacaagtctt gcttttcgaa catgggtataa tagccttggt 900
tgtgttagca aagtggaaac tatcagcatt gttgaaatgc ttaagactgc tgctgataat 960
tttgaatat aagttttgaa atctaaatgt caattttcta caaattataa aaataaaactc 1020
cactctctat gctaaaaaaa aaaaaaagga attc 1054

```

<210> 112

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003090

<400> 112

taatagcctt gtttgtgtta gcaaagtgga atctatcagc attgttgaaa tgcttaagac 60

<210> 113

<211> 2033

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003158

<400> 113

```

gaattccggg actgagctct tgaagacttg ggtccttggt cgcaggtgga ggcacgggtc 60
tcactccatt gccagggcca gagtgcggga ttttggataa gaaacttcag tgaaggccgg 120
gcgcggtgct catgcccgtat atcccagcat tttcggaggc cgaggcatca tggaccgatc 180
taaagaaaac tgcatttcag gacctgttaa ggctacagct ccagttggag gtccaaaacg 240
tggtctcgtg actcagcaat ttcttgttca gaatccatta cctgtaaaata gtggccaggc 300
tcagcgggtc ttgtgtcctt caaattcttc ccagcgcgtt cctttgcaag cacaaaagct 360
tgtctccagt cacaagccgg ttcagaatca gaagcagaag caattgcagg caaccagtgt 420
acctatcct gtctccaggc cactgaataa caccctaaaag agcaagcagc ccctgccatc 480
gcacctgaaa ataactctga ggaggaaact gcatcaaaaac agaaaaatga agaatacaaa 540
agaggcagtg gctttggaag actttgaaat tgggtgccct ctgggtaaaag gaaagtgttg 600
taatgtttat ttggcaagag aaaagcaaaag caagtttatt ctgggtctta aagtgttatt 660
taaagctcag ctggagaaaag ccggagtgga gcatcagctc agaagagaag tagaaataca 720
gtcccacctt cggcatccta atattcttag actgtatggt tatttccatg atgctaccag 780
agtctaccta attctggaat atgcaccact tggaaacagtt tatagagaac ttcagaaact 840
ttcaaaagtt gatgagcaga gaactgctaa cttatataac agaattgcaa atgccctgtc 900
ttactgtcat tcgaagagag ttattcatag agacattaag ccagagaact tacttcttgg 960
atcagctgga gagcttaaaa ttgcagattt tgggtgggtc gtacatgctc catcttccag 1020
gaggaccact ctctgtggca ccctggacta cctgccccct gaaatgattg aaggtcggat 1080
gcatgatgag aaggtggatc tctggagcct tggagtctct tgctatgaat ttttagttgg 1140
gaagcctcct tttgaggcaa acacatacca agagacctac aaaagaatat cacgggttga 1200
attcacattc cctgactttg taacagaggg agccagggac ctcatctcaa gactgttgaa 1260
gcataatccc agccagaggc caatgctcag agaagtactt gaacacctct ggatcacagc 1320
aaattcatca aaacctcaa attgccaaaa caaagaatca gctagcaaac agtcttagga 1380
atcgtgcagg gggagaaaac cttgagccag ggctgccata taacctgaca ggaacatgct 1440
actgaagttt attttaccat tgactgtctc cctcaatcta gaacgtaca caagaaatat 1500
tttgttttta ctacgaggt gtgccttaac ctccctattc agaaagctcc acatcaataa 1560
acatgacact ctgaagttaa agtagccacg agaattgtgc tacttatact ggaacataat 1620
ctggaggcaa gggtcagctg cagtcgaacc ttgcctccag attatgaacc agtataagta 1680
gcacaattct cgtggctact ttcacttcag agtgtcatgt ttattgatgt ggagctttct 1740
gaatagggag gtttaaggac acctgctgag taaaacaaat atttcttgtg tagcgttctt 1800
aggaatctgg tgtctgtccg gcccggtag gctgttggg tttctagtcc tcttaccat 1860
catctccata tgagagtgtg aaaataggaa cacgtgctct acctccattt agggatttgc 1920
ttgggataca gaagaggcca tgtgtctcag agctgttaag ggcttatttt tttaaaacat 1980
tgagtcata gcatgtgtgt aaactttaaa tatgcaggcc ttcgtggctc gag 2033

```

<210> 114

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003158

<400> 114

ttgggtttct agtcctcctt accatcatct ccatatgaga gtgtgaaaat aggaacacgt 60

<210> 115
 <211> 1421
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003258

<400> 115
 acttactgcg ggacggcctt ggagagtact cgggttcgtg aacttcccgg aggcgcaatg 60
 agctgcatta acctgcccac tgtgtgtccc ggctcccca gcaagaccg ggggcagatc 120
 caggtgattc tcgggccgat gttctcagga aaaagcacag agttgatgag acgcgtccgt 180
 cgcttcagga ttgctcagta caagtgcctg gtgatcaagt atgocaaaga cactcgctac 240
 agcagcagct tctgcacaca tgaccggaac accatggagg cgctgcccgc ctgcctgctc 300
 cgagacgtgg cccaggaggc cctgggctgt gctgtcatag gcatcgacga ggggcagttt 360
 ttccctgaca tcatggagtt ctgagaggcc atggccaacg ccgggaagac cgtaattgtg 420
 gctgactgg atgggacctt ccagaggaa ccatctgggg ccatcctgaa cctggtgccc 480
 ctggccgaga gcgtggtgaa gctgacggcg gtgtgcatgg agtgcttccg ggaagccgcc 540
 tataccaaga ggctcggcac agagaaggag gtgaggtga ttgggggagc agacaagtac 600
 cactccgtgt gtcggctctg ctacttcaag aaggcctcag gccagcctgc cgggccggac 660
 aacaaagaga actgccaggt gccaggaaag ccaggggaa cctgggctgc cagggaagctc 720
 tttgcccac agcagattct gcaatgcagc cctgccaaact gagggacctg caagggccgc 780
 ccgtccctt cctgccactg ccgcctactg gacgtgccc tgcattgctc ccagccactc 840
 caggaggaag tcgggaggcg tggagggtga ccacacctt gccttctggg aactctcctt 900
 tgtgtggctg cccacactgc cgcattgctc ctctctctc acccaactgt ctgcttaaaag 960
 ctccctctc agctgtgtgg acgatcgccc aggtgtggagc tggcccgcgt tgggtggcctg 1020
 ggatctggca cactccctct ccttgggggt agggacagag cccacgctg ttgacatcag 1080
 cctgttctt cccctctgag gctttcactg ctgagtttct gttctccctg ggaagcctgt 1140
 gccagcact ttgagcctt gccacactg aggttaggc ctctctgcct gggatgggct 1200
 cccaccctc cctgaggatg gcctggattc acgcctctt gtttcctttt gggctcaaaag 1260
 cccttctac ctctggtgat ggtttccaca ggaacaacag catctttcac caagatgggt 1320
 ggacccaacc ttgctgggac ttggatccca ggggcttatc tcttcaagtg tggagagggc 1380
 aggttccacg cctctgctgt agcttatgaa attaactaat t 1421

<210> 116
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003258

<400> 116
 cttctacct ctggtgatgg tttccacagg aacaacagca tctttcacca agatgggtgg 60

<210> 117
 <211> 913
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003311

<400> 117
 agagccggcg ccgtcaccgc ccgcattgcc gctcccagtc ccgcgctcgg caccacatga 60
 aatccccga cgaggtgcta cgcgaggcg agttggagaa gcgcagcgac agcctcttcc 120
 agctatggaa gaagaagcgc ggggtgctca cctcagaccg cctgagcctg ttccccgcca 180
 gccccgcgc gcgcccgaag gagctgctgt tccactccat cctcaagggt gactgctgtg 240
 agcgcacggg caagtacgtg tacttcacca tcgtcaccac cgaccacaag gagatcgact 300
 tccgtgcgc gggcgagagc tgetggaacg cggccatcgc gctggcgctc atcgatttcc 360
 agaaccgcgc cgcctgcag gactttcgca gccgccagga acgcaccgca cccgccgcac 420

```

ccgccgagga cgccgtggct gccgcggcgc ccgcaccctc cgagccctcg gagccctcca 480
ggccatcccc gcagcccaaa cccgcgacgc catgagcccg ccgcgggcca tacgtggac 540
gagtcggacc gaggctagga cgtggccggc gctctccagc cctgcagcag aagaacttcc 600
cgtgcgcgcg gatcctcgct ccgttgacgc ggccgcttaa gttattggac tatctaatat 660
ctatgtatatt atttcgctgg ttctttgtag tcacatattt tatagtctta atatcttggt 720
tttgcatacac tgtgccatt gcaataaat cacttgacca gtttgctttt ctaccatccg 780
gctgtggctc agtgagactc ctgctgggag ggtggaggcc caggaatggg cgggcaggac 840
accctcatcc agtcctgcgg ggctggtgtg aaaggcgtg ggaaccggct ttgaatgaat 900
aatgaatcg tgt 913

```

<210> 118
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003311

<400> 118
 atttcgctgg ttctttgtag tcacatattt tatagtctta atatcttggt tttgcatacac 60

<210> 119
 <211> 1723
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003376

<400> 119
 tcgcggaggg ttggggcagc cgggtagctc ggaggctcgt gcgctggggg ctagcaccag 60
 cgctctgtcg ggaggcgagc cggttaggtg gaccggtcag cggactcacc ggccagggcg 120
 ctcggtgctg gaatttgata ttattgata cgggttttat ccctcttctt tttctttaa 180
 catTTTTTTT taaaactgta ttgtttctcg ttttaattta tttttgcttg ccattcccca 240
 ctggaatcgg gccgacggct tggggagatt gctctacttc cccaaatcac tgtggatttt 300
 ggaaaccagc agaaagagga aagaggtagc aagagctcca gagagaagtc gaggaagaga 360
 gagacggggg cagagagagc gcgcgggcgt gcgagcagcg aaagcgacag gggcaaagt 420
 agtgacctgc ttttgggggt gaccgcggga gcgcggcgtg agccctcccc cttgggatcc 480
 cgacgtgac cagtcgcgtg gacggacaga cagacagaca ccgccccag cccagctac 540
 cacctcctcc ccggccggcg gcggacagtg gacgcggcgg cgagccgcgg gcaggggcgg 600
 gagccgcgc ccggaggcgg ggtggagggg gtcggggctc gcggcgctcg actgaaactt 660
 ttctgccaac ttctgggctg ttctcgcttc ggaggagccg tggtcgcgcg gggggaagcc 720
 gagccgagcg gagccgcgag aagtgtatgc tcgggcccgg aggagccgca gccggaggag 780
 ggggaggagg aagaagagaa ggaagaggag agggggccgc agtggcgact cggcgctcgg 840
 aagccggggt catggacggg tgaggcggcg gtgtgcgcag acagtgtctc agccgcgcgc 900
 gtcctccagg ccctggcccg ggctcgggc cggggaggaa gtagtagctc ccgaggcgcc 960
 gaggagagcg ggccgcccc cagcccgagc cggagaggga gcgcgagccg ccgaggcccc 1020
 ggtcgggcct ccgaaaccat gaactttctg ctgtcttggt tgcattggag ccttgccctt 1080
 ctgctctacc tccaccatgc caagtggctc caggctgcac ccatggcaga aggaggagg 1140
 cagaatcatc acgaagtggg gaagtctatg gatgtctatc agcgagcta ctgccatcca 1200
 atcgagaccc tgggtggacat cttccaggag taccctgatg agatcgagta catcttcaag 1260
 ccactcctgt tgccctgat gcgatgcggg ggtgctgca atgacaggg cctggagtgt 1320
 gtgcccactg aggcgtccaa catcaccatg cagattatgc ggatcaaacc tcaccaaggc 1380
 cagcacatag gagagatgag ctctctacag cacaacaaat gtgaatgcag accaaagaaa 1440
 gatagagcaa gacaagaaaa aaaatcagtt cgaggaaagg gaaaggggca aaaacgaaa 1500
 cgcaagaaat cccggtataa gtccctggag gttccctgtg ggcccttgctc agagcggaga 1560
 aagcatttgt ttgtacaaga tccgcagacg tgtaaatgtt cctgcaaaaa cacagactcg 1620
 cgttgcaagg cgaggcagct tgagttaaac gaacgtactt gcagatgtga caagccgagg 1680
 cggtagagcc ggacaggagga aggagcctcc ctcagggttt cg 1723

<210> 120
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003376

<400> 120
 ccagcacata ggagagatga gcttctctaca gcacaacaaa tgtgaatgca gaccaaaaga 60

<210> 121
 <211> 2834
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003406

<400> 121
 gccactccc accgccagct ggaaccctgg ggactacgac gtccctcaaa ccttgcttct 60
 aggagataaa aagaacatcc agtcatggat aaaaatgagc tggttcagaa ggccaaactg 120
 gccgagcagg ctgagcgata tgatgacatg gcagcctgca tgaagtctgt aactgagcaa 180
 ggagctgaat tatccaatga ggagaggaat cttctctcag ttgcttataa aaatgttgta 240
 ggagcccgta ggtcatcttg gagggctgct tcaagtattg aacaaaagac ggaagggtgct 300
 gagaaaaaac agcagatggc tcgagaatac agagagaaaa ttgagacgga gctaagagat 360
 atctgcaatg atgtactgtc tcttttggaa aagttcttga tccccaatgc ttcacaagca 420
 gagagcaaaag tcttctatctt gaaaatgaaa ggagattact accgttactt ggctgagggt 480
 gccgctggtg atgacaagaa agggattgtc gatcagtcac aacaagcata ccaagaagct 540
 tttgaaatca gcaaaaagga aatgcaacca acacatccta tcagactggg tctggccctt 600
 aacttctctg tgttctatta tgagattctg aactccccag agaaagcctg ctctcttgca 660
 aagacagctt ttgatgaagc cattgctgaa cttgatacat taagtgaaga gtcatacaaa 720
 gacagcacgc taataatgca attactgaga gacaacttga cattgtggac atcggatacc 780
 caaggagacg aagctgaagc aggagaagga ggggaaaatt aaccggcctt ccaacttttg 840
 tctgcctcat tctaaaattt acacagtaga ccatttgtca tccatgctgt cccacaaata 900
 gttttttgtt tacgatttat gacaggttta tgttacttct atttgaattt ctatatttcc 960
 catgtggttt ttatgtttaa tattagggga gtagagccag ttaacattta gggagttatc 1020
 tgttttcate ttgaggtggc caatatgggg atgtggaatt ttatacaag ttataagtgt 1080
 ttggcatagt acttttggta cattgtggct tcaaaagggc cagtgtaaaa ctgcttccat 1140

gtctaagcaa agaaaactgc ctacatactg gtttgtcctg gcgggggaata aaagggatca 1200
 ttggttccag tcacaggtgt agtaattgtg ggtactttta ggtttggagc acttacaagg 1260
 ctgtggtaga atcatacccc atggatacca catattaaac catgtatatc tgtggaatac 1320
 tcaatgtgta cacctttgac tacagctgca gaagtgttcc tttagacaaa gttgtgaccc 1380
 attttactct ggataagggc agaaacggtt cacattccat tatttgtaaa gttacctgct 1440
 gttagctttc attatttttg ctacactcat tttatttgta tttaaatgtt ttaggcaacc 1500
 taagaacaaa tgtaaaagta aagatgcagg aaaaatgaat tgcttgggtat tcattacttc 1560
 atgtatatca agcacagcag taaaacaaaa acccatgtat ttaacttttt tttaggattt 1620
 ttgcttttgt gatttttttt tttttttttt gatacttgcc taacatgcat gtgctgtaaa 1680
 aatagttaac agggaaataa cttgagatga tggctagctt tgtttaatgt cttatgaaat 1740
 tttcatgaac aatccaagca taattgttaa gaacacgtgt attaaattca tgtaagtgga 1800
 ataaaagttt tatgaatgga cttttcaact actttctcta cagcttttca tgtaaattag 1860
 tcttggttct gaaacttctc taaaggaaat tgtacatttt ttgaaattta ttccttattc 1920
 cctcttgcca gctaattggc tcttaccag tttaaacaca aaatttatca taacaaaaat 1980
 actactaata taactactgt ttccatgtcc catgatcccc tctcttctc cccaccctga 2040
 aaaaaatgag ttcctatttt ttctgggaga gggggggatt gattagaaaa aaatgtagt 2100
 tgttccattt aaaatttttg catatggcat tttctaactt aggaagccac aatgttcttg 2160
 gcccatcatg acattgggta gcatttaact taagttttgt gcttccaaat cacttttttg 2220
 tttttaagaa tttcttgata ctcttatagc ctgccttcaa ttttgatcct ttattctttc 2280
 tatttgtcag gtgcacaaga ttaccttctt gtttttagcct tctgtcttgt caccaacct 2340
 tcttacttg tggccatgta cttggaaaaa ggccgcatga tctttctggc tccactcagt 2400

```

gtctaaggca ccttgcttcc tttgcttgca toccacagac tatttccctc atcctattta 2460
ctgcagcaaa tctctcctta gttgatgaga ctgtgtttat ctccctttaa aaccctacct 2520
atcctgaatg gtctgtcatt gtctgccttt aaaatccttc ctctttcttc ctccctctatt 2580
ctctaaataa tgatggggct aagttatacc caaagctcac tttacaaaat atttcctcag 2640
tactttgcag aaaacaccaa acaaaaatgc cattttaaaa aaggtgtatt ttttctttta 2700
gaatgtaagc tcctcaagag cagggacaat gttttctgta tgttctattg tgcctagtac 2760
actgtaaatg ctcaataaat attgatgatg ggaggcagtg agtcttgatg ataaggggtga 2820
gaaactgaaa tccc 2834

```

```

<210> 122
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_003406

```

```

<400> 122
tttagccttc tgtcttgtca ccaaccattc ttacttgggtg gccatgtact tggaaaaagg 60

```

```

<210> 123
<211> 1938
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_003504

```

```

<400> 123
gatttggcgg gagtcttgac cgcgcgcggg ctcttggtac ctccagcgga ggcgccaggcg 60
tccggccgcc gtggctatgt tctgtgtcga tttccgcaaa gagttctacg aggtggtcca 120
gagccagagg gtccttctct tctgtgcctc ggacgtggat gctctgtgtg cgtgcaagat 180
ccttcaggcc ttgttccagt gtgaccacgt gcaatatacg ctggttccag tttctgggtg 240
gcaagaactt gaaactgcat ttcttgagca taaagaacag tttcattatt ttattctcat 300
aaactgtgga gctaattgtg acctattgga tattcttcaa cctgatgaag aactatatt 360
ctttgtgtgt gacacccata ggccagtcaa tgtcgtcaat gtatacaacg ataccagat 420
caaattactc attaaacaag atgatgacct tgaagttccc gcctatgaag acatcttcag 480
ggatgaagag gaggatgaag agcattcagg aaatgacagt gatgggtcag agccttctga 540
gaagcgcaca cggttagaag aggagatagt ggagcaaac atgcggagga ggcagcggcg 600
agagtgggag gcccgagaa gagacatcct ctttgactac gagcagtatg aatatcatgg 660
gacatcgtca gccatggtga tgtttgagct ggcttggtat ctgtccaagg acctgaatga 720
catgctgtgg tgggccatcg ttggactaac agaccagtgg gtgcaagaca agatcactca 780
aatgaaatac gtgactgatg ttgggtgtcct gcagcgccac gtttcccgcc acaaccaccg 840
gaacgaggat gaggagaaca cactctccgt ggactgcaca cggatctcct ttgagtatga 900
cctccgcctg gtgctctacc agcactggtc cctccatgac agcctgtgca acaccagcta 960
taccgcagcc aggttcaagc tgtggtctgt ccatggacag aagcggctcc aggagtctct 1020
tgcagacatg ggtcttcccc tgaagcaggt gaagcagaag ttccaggcca tggacatctc 1080
cttgaaggag aatttgcggg aaatgattga agagtctgca aataaatttg ggatgaagga 1140
catgcgcgtg cagactttca gcattcattt tgggttcaag cacaagtttc tggccagcga 1200
cgtggtcttt gccaccatgt ctttgatgga gagccccgag aaggatggct cagggacaga 1260
tcacttcac caggctcttg acagcctctc caggagtaac ctggacaagc tgtaccatgg 1320
cctggaactc gccaaagc agctgcgagc caccagcag accattgcca gctgcctttg 1380
caccaacctc gtcactcctc aggggccttt cctgtactgc tctctcatgg agggcactcc 1440
agatgtcatg ctgttctcta ggccggcatc cctaagcctg ctccagcaaac acctgctcaa 1500
gtcctttgtg tgttcgacaa agaaccggcg ctgcaaaactg ctgcccctgg tgatggctgc 1560
ccccctgagc atggagcatg gcacagtgac cgtggtgggc atccccccag agaccgacag 1620
ctcggacagg aagaactttt ttgggagggc gtttgagaag gcagcggaaa gcaccagctc 1680
cgggatgctg cacaaccatt ttgacctctc agtaattgag ctgaaagctg aggatcggag 1740
caagtttctg agcgactta tttccctcct gtcctaggaa tttgattctt ccagaatgac 1800
cttcttattg atgtaactgg ctttcattta tattgtaagt tatggacatg atttgagatg 1860
tagaagccat tttttattaa ataaaatgct tatttttaggc tccgtcccca aaaaaaaaaa 1920

```

aaaaaaaaa aaaaaaaaa 1938

<210> 124

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003504

<400> 124

caagtttctg gacgcaactta tttccctcct gtccctaggaa tttgattctt ccagaatgac 60

<210> 125

<211> 2346

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003600

<400> 125

acaaggcagc	ctcgctcgag	cgcaggccaa	tcggctttct	agctagaggg	tttaactcct	60
atttaaaaag	aagaaccttt	gaattctaac	ggctgagctc	ttggaagact	tgggtccttg	120
ggtcgcaggt	gggagccgac	gggtgggtag	accgtggggg	atatctcagt	ggcggacgag	180
gacggcgggg	acaaggggcg	gctggtcgga	gtggcggagc	gtcaagtccc	ctgtcggttc	240
ctccgtccct	gagtgtcctt	ggcgctgcct	tgtgcccgcc	cagcgccttt	gcatacgcctc	300
ctgggcaccg	aggcgccttg	taggatactg	cttggttactt	attacagcta	gagggcatcat	360
ggaccgatct	aaagaaaact	gcatttcagg	acctgttaag	gctacagctc	cagttggagg	420
tccaaaacgt	gttctcgtga	ctcagcaatt	tccttgtcag	aatccattac	ctgtaaataag	480
tggccaggct	cagcgggtct	tgtgtccttc	aaattcttcc	cagcgcattc	ctttgcaagc	540
acaaaagctt	gtctccagtc	acaagccggt	tcagaatcag	aagcagaagc	aattgcaggc	600
aaccagtgtg	cctcatcctg	tctccaggcc	actgaataac	acccaaaaga	gcaagcagcc	660
cctgccatcg	gcacctgaaa	ataatcctga	ggaggaactg	gcatacaaac	agaaaaatga	720
agaatcaaaa	aagaggcagt	gggctttgga	agactttgaa	attgggtcgcc	ctctgggttaa	780
aggaaaagttt	ggtaatgttt	atttggcaag	agaaaagcaa	agcaagttta	ttctggctct	840
taaagtgtta	tttaaagctc	agctggagaa	agccggagtg	gagcatcagc	tcagaagaga	900
agtagaaaata	cagtcaccac	ttcggcatcc	taatttctt	agactgtatg	gttattttcca	960
tgatgctaacc	agagtctaacc	taattctgga	atatgcacca	cttggaaacag	tttatagaga	1020
acttcagaaa	ctttcaaagt	ttgatgagca	gagaactgct	acttatataa	cagaattggc	1080
aaatgccctg	tcttactgtc	attcgaagag	agttattcat	agagacatta	agccagagaa	1140
cttacttctt	ggatcagctg	gagagcttaa	aattgcagat	tttgggtggg	cagtacatgc	1200
tccatcttcc	aggaggacca	ctctctgtgg	cacctgggac	tacctgcccc	ctgaaatgat	1260
tgaaggtcgg	atgcatgatg	agaaggtgga	tctctggagc	cttggagtgc	tttgctatga	1320
atttttagtt	gggaagcctc	cttttgaggc	aaacacatac	caagagacct	acaaaagaat	1380
atcacggggt	gaattcacat	tccctgactt	tgtaacagag	ggagccaggg	acctcatttc	1440
aagactgttg	aagcataatc	ccagccagag	gccaatgctc	agagaagtac	ttgaacaccc	1500
ctggatcaca	gcaaattcat	caaaaccatc	aaattgccaa	aacaaagaat	cagctagcaa	1560
acagtcttag	gaatcgtgca	gggggagaaa	tccttgagcc	agggctgcca	tataacctga	1620
caggaacatg	ctactgaagt	ttattttacc	attgactgct	gccctcaatc	tagaacgcta	1680
cacaagaaat	atttgtttta	ctcagcaggt	gtgccttaac	ctccctattc	agaaagctcc	1740
acatcaataa	acatgacact	ctgaagtga	agtagccacg	agaattgtgc	tacttatact	1800
ggttcataat	ctggaggcaa	ggttcgactg	cagccgcccc	gtcagcctgt	gctaggcatg	1860
gtgtcttcac	aggaggcaaa	tccagagcct	ggctgtgggg	aaagtgacca	ctctgccctg	1920
accccgatca	gttaaggagc	tgtgcaataa	ccttcctagt	acctgagtga	gtgtgtaact	1980
tattgggttg	gcgaagcctg	gtaaagctgt	tgggaatgagt	atgtgattct	ttttaagtat	2040
gaaaataaag	atatatgtac	agacttgtat	tttttctctg	gtggcattcc	tttaggaatg	2100
ctgtgtgtct	gtccggcacc	cgggtaggcc	tgattgggtt	tctagtccctc	cttaaccact	2160
tatctcccat	atgagagtgt	gaaaaatagg	aacacgtgct	ctacctccat	ttagggtatt	2220
gcttgggata	cagaagaggc	catgtgtctc	agagctgtta	agggcttatt	tttttaaaac	2280

attggagtca tagcatgtgt gtaaaacttta aatatgcaaa taaataagta tctatgtcta 2340
aaaaaa 2346

<210> 126
<211> 60
<212> DNA
<213> Homo sapiens

<300>

<308> NM_003600

<400> 126
agagtgtgaa aaataggaac acgtgctcta cctccattta gggatttgct tgggatacag 60

<210> 127
<211> 853
<212> DNA
<213> Homo sapiens

<300>

<308> NM_003641

<400> 127
ctagtcctga cttcacttct gatgaggaag cctctctcct tagccttcag cctttcctcc 60
caccctgcc aagtaattt gatcctcaag aagttaaacc acacctcatt ggtccctggc 120
taattcacca atttacaac agcaggaaat agaaacttaa gagaaatata cacttctgag 180
aaactgaaac gacaggggaa aggaggtctc actgagcacc gtcccagcat ccggacacca 240
cagcggccct tcgtccacg cagaaaacca cacttctcaa accttcactc aacacttcct 300
tccccaaagc cagaagatgc acaaggagga acatgagggt gctgtgctgg gggcaccccc 360
cagcaccatc cttccaaggt ccaccgtgat caacatccac agcgagacct ccgtgcccga 420
ccatgtcgtc tgggtccctgt tcaacaccct cttcttgaac tgggtgctgc tgggcttcat 480
agcattcgcc tactcogtga agtctagggg caggaagatg gttggcgacg tgaccggggc 540
ccaggcctat gcctccacg ccaagtgcct gaacatctgg gcctgatcc tgggcattcc 600
catgaccatt ggattcatcc tgtcactggt attcggtct gtgacagtct accatattat 660
gttacagata atacaggaaa aacgggggtta ctagtagccg cccatagcct gcaacctttg 720
cactccactg tgcaatgctg gccctgcacg ctggggctgt tgcccctgcc cccttggtcc 780
tgcccctaga tacagcagtt tataaccaca cactgtcta cagtgtcatt caataaagtg 840
cacgtgcttg tga 853

<210> 128
<211> 60
<212> DNA
<213> Homo sapiens

<300>

<308> NM_003641

<400> 128
attatgttac agataatata ggaaaaacgg ggttactagt agccgcccac agcctgcaac 60

<210> 129
<211> 1280
<212> DNA
<213> Homo sapiens

<300>

<308> NM_003756

<400> 129


```

gaaagatggc gtcccgaag gaaggtaccg gctctactgc cacctcttcc agctccaccg 60
ccggcgagc agggaaagg aaaggcaaaag gcggctcggg agattcagcc gtgaagcaag 120
tgcagataga tggccttggt gtattaaaga taatcaaaca ttatcaagaa gaaggacaag 180
gaactgaagt tgttcaagga gtgcttttgg gtctggttgt agaagatcgg cttgaaatta 240
ccaactgctt tcctttccct cagcacacag aggatgatgc tgactttgat gaagtccaat 300
atcagatgga aatgatgcgg agccttcgcc atgtaaacad tgatcatctt cacgtgggct 360
ggatcagtc cacatactat ggctcattcg ttaccggggc actcctggac tctcagttta 420
gttaccagca tgccattgaa gaatctgtcg ttctcattta tgatcccata aaaactgccc 480
aaggatctct ctcactaaag gcatacagac tgactcctaa actgatggaa gtttgtaaag 540
aaaaggattt ttccctgaa gcattgaaaa aagcaaatat caccttgag tacatgtttg 600
aagaagtgcc gattgtaatt aaaaattcac atctgatcaa tgtcctaatt tgggaacttg 660
aaaagaagtc agctgttgca gataaacatg aattgctcag ccttgccagc agcaatcatt 720
tggggaagaa tctacagttg ctgatggaca gagtggatga aatgagccaa gatatagtta 780
aatacaacac atacatgagg aatactagta aacaacagca gcagaaacat cagtatcagc 840
agcgtcgcca gcaggagaat atgcagcgcc agagccgagg agaaccctcg ctccctgagg 900
aggacctgtc caaactcttc aaaccaccac agccgcctgc caggatggac tcgctgctca 960
ttgcaggcca gataaacact tactgccaga acatcaagga gttcactgcc caaaacttag 1020
gcaagctctt catggcccag gctcttcaag aatacaacaa ctaagaaaag gaagtttcca 1080
gaaaagaagt taacatgaac tcttgaagtc acaccagggc aactcttgga agaaatatat 1140
ttgcatattg aaaagcacag aggatttctt tagtgtcatt gccgattttg gctataacag 1200
tgtctttcta gccataataa aataaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1260
aaaaaaaaa aaaaaaaaaa 1280

```

<210> 130

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003756

<400> 130

```

tgagccaaga tatagttaaa tacaacacat acatgaggaa tactagtaaa caacagcagc 60

```

<210> 131

<211> 839

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003832

<400> 131

```

aagccacagg ctccctggct ggcgtcagct aaagtggctg ttgggtgtcc gcaggcttct 60
gcctggccgc cgccgcctat aagctaccag gaggagcttt acgacttccc gtccctgctg 120
aagtggcggg cagcatcgca aggtagcgca gaagcttctc aatggccagc gccagctgca 180
gccccggcgg cgactcgcc tcacctgagc ctgggaggaa aattcttcca aggatgatct 240
cccactcaga gctgaggaag cttttctact cagcagatgc tgtgtgtttt gatgttgaca 300
gcacggtcat cagtgaagaa ggaatcggat gctttcattg gatttggagg aaatgtgatc 360
aggcaacaag tcaaggataa cgccaaatgg tatatcactg attttgtaga gctgctggga 420
gaaccggaag aataacatcc attgtcatac agctccaaac aacttcagat gaatttttac 480
aagttacaca gattgatact gtttgcttac aattgcctat tacaacttgc tataaaaagt 540
tggtagagat gatctgcact gtcaagtaaa ctacagttag gaatcctcaa agattggttt 600
gtttgttttt aactgtagtt ccagttattat atgatcacta tcgatttctt ggagagtttt 660
gtaatctgaa ttctttatgt atattcctag ctatatttca taaaagtgt ttttaagagt 720
gagagtcaat taaacacctt tactcttagg aatatagatt cggcagcctt cagtgaatat 780
tgggtttttt ccctttggta tgtcaataaa agtttatcca tgtgtcagaa aaaaaaaaaa 839

```

<210> 132

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003832

<400> 132

gaagaaggaa tcggatgctt tcattggatt tggaggaaat gtgatcaggc aacaagtcaa 60

<210> 133

<211> 3128

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003981

<400> 133

gcttcgcccc	gtggcgcggt	ttgaaat	ttt	gcggggctca	acggctcgcg	gagcggctac	60
gcggaagtac	atcgccggtg	tttgcgggtg	gttgttgctc	tcggggccgt	gtggagtagg		120
tctggacctg	gactcacggc	tgcttggagc	gtccgccatg	aggagaagtg	aggtgctggc		180
ggaggagtcc	atagtatgtc	tgcaaaaagc	cctaaatcac	cttcgggaaa	tatgggagct		240
aattgggatt	ccagaggacc	agcggttaca	aagaactgag	gtggtaaaga	agcatatcaa		300
ggaactcctg	gatatgtaga	ttgctgaaga	ggaaagcctg	aaggaaagac	tcatcaaaag		360
catatccgtc	tgctcagaaag	agctgaacac	tctgtgcagc	gagttacatg	ttgagccatt		420
tcaggaagaa	ggagagacga	ccatcttgca	actagaaaaa	gattttgcgc	cccaagtggg		480
attgatgcga	aaacagaaaa	aggagagaaa	acaggaactg	aagctacttc	aagagcaaga		540
tcaagaactg	tgcgaaat	tttgtatgcc	ccactatgat	attgacagtg	cctcagtgcc		600
cagcttagaa	gagctgaacc	agttcaggca	acatgtgaca	acttttaggg	aaacaaaggc		660
ttctaggcgt	gaggagtttg	tcagtataaa	gagacagatc	atactgtgta	tggaagaatt		720
agaccacacc	ccagacacaa	gctttgaaag	tgtgtggtg	tgtgaagacg	aagatgcctt		780
ttgtttgtct	ttggagaata	ttgcaacact	acaaaagttg	ctacggcagc	tggaaatgca		840
gaaatcacaa	aatgaagcag	tgtgtgaggg	gctgcgtact	caaatccgag	agctctggga		900
caggttgcaa	atacctgaag	aagaaagaga	agctgtggcc	accattatgt	ctgggtcaaa		960
ggccaagggtc	cggaagcgc	tgcaattaga	agtggatcgg	ttggaagaac	tgaaaatgca		1020
aaacatgaag	aaagtgattg	aggcaattcg	agtggagctg	gttcagtagt	gggaccagtg		1080
cttttatagc	caggagcaga	gacaagcttt	tgcccccttc	tgtgctgagg	actacacaga		1140
aagtctgctc	cagctccacg	atgctgagat	tgtgctggta	aaaaactact	atgaagttca		1200
caaggaactc	tttgaagggtg	tcagaaagt	ggaagaaacc	tggaggcttt	tcttagagtt		1260
tgagagaaaa	gcttcagatc	caaatcgatt	tacaaaccga	ggaggaaatc	ttctaaaaga		1320
agaaaaacaa	cgagccaagc	tccagaaaat	gctgcccaag	ctggaagaag	agttgaaggc		1380
acgaattgaa	ttgtgggaac	aggaacattc	aaaggcattt	atggtgaatg	ggcagaaatt		1440
catggagtat	gtggcagaac	aatgggagat	gcacgattg	gagaaagaga	gagccaagca		1500
ggaaagacaa	ctgaagaaca	aaaaacagag	agagacagag	atgctgtatg	gcagcgctcc		1560
tcgaacacct	agcaagcggc	gaggactggc	tcccaataca	ccgggcaaag	cacgtaagct		1620
gaacactacc	accatgtcca	atgctacggc	caatagtagc	attcggccta	tctttggagg		1680
gacagtctac	cactcccccg	tgtctcgact	tcctccttct	ggcagcaagc	cagtcgctgc		1740
ttccacctgt	tcagggaaga	aaacaccccg	tactggcagg	catggagcca	acaaggagaa		1800
cctggagctc	aacggcagca	tcctgagtgg	tgggtaccct	ggctcggccc	ccctccagcg		1860
caacttcagc	attaattctg	ttgccagcac	ctattctgag	tttgccaagg	atccgtccct		1920
ctctgacagt	tccactgttg	ggcttcagcg	agaactttca	aaggcttcca	aatctgatgc		1980
tacttctgga	atcctcaatt	caaccaacat	ccagtcctga	gaagccctga	tcagtcaacc		2040
agctgtggct	tcctgtgcct	agactggacc	taattatatg	ggggtgactt	tagtttttct		2100
tcagcttagg	cgtgcttgaa	accttggcca	ggttccatga	ccatgggcct	aacttaaaga		2160
tgtgaatgag	tggttacagtt	gaaagcccat	cataggttta	gtggtcctag	gagacttggt		2220
tttgacttat	atacatgaaa	agtttatggc	aagaagtgca	aatttttagca	tatggggcct		2280
gacttctcta	ccacataatt	ctacttgctg	aagcatgac	aaagcttggt	ttatctcacc		2340
actgtagtaa	aatgattgac	tatgccatc	cctgggggta	atgttggcat	gtatacctgt		2400
aactagtaat	taacatcttt	tttgtttatg	catgttcaat	taatgctgta	gctatcatag		2460
ctttgctctt	acctgaagcc	ttgtccccc	cacacaggac	agccttcctc	ctgaagagaa		2520

```

tgtctttgtg tgtccgaagt tgagatggcc tgccctactg ccaaagaggt gacaggaagg 2580
ctgggagcag ctttgttaaa ttgtgttcag ttctgttaca cagtgcattg ccctttgttg 2640
ggggatgca tgtatgaaca cacatgcttg tcggaacgct ttctcggcgt ttgtcccttg 2700
gctctcatct ccccatctcc tgtgcctact ttgcctgagt tcttctaccc ccgcagttgc 2760
cagccacatt gggagtctgt ttgttccaat gggttgagct gtctttgtcg tggagatctg 2820
gaactttgca catgtcacta ctggggaggt gttcctgctc tagcttccac gatgaggcgc 2880
cctctttacc tatcctctca atcactactc ttcttgaagc actattattht attcttccgc 2940
tgtctgcctg cagcagtact actgtcaaca tagtgtaaat ggttctcaaa agcttaccag 3000
tgtggacttg gtgttagcca cgctgtttac tcatacagta cgtgtcctgt ttttaaaata 3060
tacaattatt cttaaaaata aattaaaatc tgtatactta catttcaaaa agaaaaaaa 3120
aaaaaaa 3128

```

```

<210> 134
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_003981

```

```

<400> 134
tgcagcagta ctactgtcaa catagtgtaa atggttctca aaagcttacc agtgtggact 60

```

```

<210> 135
<211> 1816
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_004029

```

```

<400> 135
ggcaccagg gtccggcctg cgccttcccc ccaggccttg acaactggttc aacacctgtg 60
acttcatgtg tgcgcgcgg ccacacctgc agtcacacct gtagccccct ctgccaaagag 120
atccataccg aggcagcgtc ggtggctaca agccctcagt ccacacctgt ggacacctgt 180
gacacctggc cacacgacct gtggcgcgg cctggcgtct gctgcgacag gagcccttac 240
ctccctgtt ataacacct accgccacct aactgcccct gcagaaggag caatggcctt 300
ggctcctgag agggcagccc cagcgtgct gtccggagag tggctccttg gagagatcag 360
cagcggctgc tatgaggggc tgcagtggct ggacgaggcc cgcacctgtt tccgcgtgcc 420
ctggaagcac ttgcgcgcga aggacctgag cgaggccgac gcgcgcacct tcaaggcctg 480
ggctgtggcc cgcggcaggt ggccgcctag cagcagggga ggtggccgc ccccgaggc 540
tgagactgcg gagcgcgcc gctggaaaac caacttccgc tgcgcactgc gcagcacgcg 600
tcgcttcgtg atgctgcggg ataactcggg ggaccgcggc gaccgcgaca aggtgtacgc 660
gctcagccgg gagctgtgct ggcgagaagg ccaggcacg gaccagactg aggcagaggc 720
ccccgcagct gtcccaccac cacagggtgg gccccaggg ccattcttgg cacacacaca 780
tgctggactc caagccccag gccccctccc tgcctcagct ggtgacaagg gggacctcct 840
gctccaggca gtgcaacaga gctgcctggc agaccatctg ctgacagcgt catggggggc 900
agatccagtc ccaaccaagg ctctggaga gggacaagaa gggcttcccc tgactggggc 960
ctgtgctgga ggcgaggccg cggccccaga gtccccgcac caggcagagc cgtacctgtc 1020
accctcccca agcgcctgca ccgcggtgca agagcccagc ccagggggcgc tggacgtgac 1080
catcatgtac aaggcccgca cgggtgctga gaaggtggtg ggacaccgca gctgcacgtt 1140
cctataccgc cccccagacc cagctgtccg ggccacagac cccagcagg tagcattccc 1200
cagccctgcc gagctcccg accagaagca gctgcgtac acggaggaa tgctgcgga 1260
cgtggccctt gggttgcacc tggagcttcg ggggccacag ctgtggggcc ggccatggg 1320
caagtgaag gtgtactgg aggtggcg agccccaggc tccgccagcc cctccacccc 1380
agcctgctgt ctgcctcgga actgtgacac ccccatcttc gacttcagag tcttcttcca 1440
agagctggtg gaattccggg cacggcagcg ccgtggctcc ccacgctata ccatttaoct 1500
gggcttcggg caggacctgt cagctgggag gcccaaggag aagagcctgt tcctggtgaa 1560
gctggaacc tggtgtgcc gagtgacct agagggcacg cagcgtgagg gtgtgtcttc 1620
cctggatagc agcagcctca gcctctgcct gtccagcgcc aacagcctct atgacgacat 1680
cgagtgtctc cttatggagc tggagcagcc cgcctagaac ccagtctaata gagaactcca 1740

```

gaaagctgga gcagcccacc tagagctggc cgcggccgcc cagtctaata aaaagaactc 1800
cagaacaaaa aaaaaa 1816

<210> 136
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_004029

<400> 136
agcagcccac ctagagctgg ccgcgccgc ccagtctaata aaaaagaact ccagaacaaa 60

<210> 137
<211> 2121
<212> DNA
<213> Homo sapiens

<300>
<308> NM_004203

<400> 137
tgaattttt ggcgcgagca gctccgcgcg cgttcacggg ccgttcccc tcacgggagt 60
cctccgccc ggcgctccga acagtcgacg gcagactccg gccgctgag ccaccgagg 120
ggtcccgctg cctccgcgga cccggaatct gggccctcgc ggaccgcgc cccgcccagt 180
cgccccaggg cttccccaca cccacggagt gaagtcagcc gcggccctgc ctgggaggaa 240
cttaccgtct accgggaaag gtggccagca gatgtgtcgg gcctggtgag aggggtgaggc 300
gagacggccc gatcgcccag ggccccgga gctgcgagg tcacccccgc ctggccttag 360
ctcagggaca ccctggattc acgtgggagc ccctgctcct gcctccccgc tcccaccact 420
gaggctggtt ggccaggcca gtcatgctag aacggcctcc tgcaactggc atgcccattg 480
ccacggaggg caccocgcca cctctgagt gcacccccat ccagtccca gcctacttcc 540
gccacgcaga acctggattc tccctcaaga ggccccagg gctcagccg agcctccac 600
ctccgcccc tgccaagggc agcattccca tcagccgcct cttccctcct cggaccccag 660
gctggcacca gctgcagccc cggcggtgt cattccgggg cgaggcctca gagaactctgc 720
agagccctgg gtatgacca agccggccag agtccttctt ccagcagagc ttccagaggc 780
tcagccgcct gggccatggc tcctacggag aggtcttcaa ggtgcgctcc aaggaggacg 840
gccggtcta tgcggtaaag cgttccatgt caccattccg gggccccaag gaccgggccc 900
gcaagtggc cgaggtggg agccacgaga aggtggggca gcacccatgc tgcgtgaggc 960
tggagggc ctgggaggag ggcgccatcc tgtacctgca gacggagctg tgcgggccc 1020
gcctgcagca aactgtgag gcctggggtg ccagcctgcc tgaggcccag gtctggggct 1080
acctgcggga cagcgtgctt gccctggccc atctgcacag ccagggcctg gtgcacctg 1140
atgtcaagcc tgccaacatc ttccctgggg cccggggccg ctgcaagctg ggtgacttcc 1200
gactgctggt ggagctgggt acagcaggag ctggtgaggt ccaggaggga gacccccgct 1260
acatggcccc cgagctgctg cagggtcctt atgggacagc agcggatgtg ttcagtctg 1320
gcctcaccat cctggaagtg gcattgcaaca tggagctgcc ccacgggtgg gagggtggtg 1380
agcagctgag ccagggtac ctgccccctg agttcactgc cggctctgtc tccgagctgc 1440
gttctgtcct tgtcatgatg ctggagccag accccaagct gcggggccac gccgaggccc 1500
tgctggcaact gcctgtgttg aggcagccgc gggcctgggg tgtgtgtgtg tgcatggcag 1560
cggaggccct gagccgaggg tgggccctgt ggcaggccct gcttgccctg ctctgctggc 1620
tctggcatgg gctggctcac cctgccagct ggctacagcc cctggggccc ccagccaccc 1680
cgcctggctc accacctgc agtttgcctc tggacagcag cctctccagc aactgggatg 1740
acgacagcct agggccttca ctctcccctg aggtgtcct ggcccgact gtggggagca 1800
cctccacccc ccggagcagg tgcaacacca gggatgccct ggacctaaat gacatcaact 1860
cagagcctcc tcggggctcc ttccctcctt ttgagcctc gaacctctc agcctgtttg 1920
aggacacct agacccaacc tgagccccag actctgcctc tgcaactttta accttttacc 1980
ctgtgtctct cccgtcgccc ttgaaagctg gggccctcgc ggaactccca tggctctctc 2040
tgccctggcg tgtctaataa aaagtatttg aaccttggga gcacccaagc ttgctcatgt 2100
ggcaaaaaaa aaaaaaaaaa a 2121

<210> 138

<211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004203

<400> 138
 ctggccgtgt ctaataaaaa gtatttgaac cttgggagca cccaagcttg ctcatgtggc 60

<210> 139
 <211> 1982
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004207

<400> 139
 ggcgagaggc gggctgaggc ggcccagcgg cggcaggtga ggcggaacca accctcctgg 60
 ccatgggagg ggcctgggtg gacgagggcc ccacaggcgt caaggcccct gacggcggct 120
 ggggctgggc cgtgctcttc ggctgtttcg tcatcactgg ctctcctac gccttcccca 180
 aggcctgtag tgtctcttc aaggagctca tacaggagtt tgggatcggc tacagcgaca 240
 cagcctggat ctctccatc ctgctggcca tgctctacgg gacaggctcg ctctgcagtg 300
 tgtgcgtgaa ccgctttggc tgccggcccg tcatgcttgt ggggggtctc ttgctgcgc 360
 tgggcatggt ggctgcgtcc ttttgccgga gcatcatcca ggtctacctc accactgggg 420
 tcatcacggg gttgggtttg gcaactcaact tccagccctc gctcatcatg ctgaaccgct 480
 acttcagcaa gcggcgcccc atggccaacg ggctggcggc agcaggtagc cctgtcttcc 540
 tgtgtgccct gagcccgctg gggcagctgc tgcaggaccg ctacggctgg cggggcggct 600
 tctcatcct gggcgccctg ctgctcaact gctgcgtgtg tgccgcactc atgaggcccc 660
 tgggtggtcag ggcccagccg ggctcggggc cgcgcgcacc ctcccggcgc ctgctagacc 720
 tgagcgtctt ccgggaccgc ggctttgtgc tttacgcctg ggccgcctcg gtcattggtgc 780
 tggggtctct cgtcccgcgc gtgttcgtgg tgagctacgc caaggacctg ggctgtcccg 840
 acaccaaggc cgccttctct ctcaccatcc tgggcttcat tgacatcttc gcgcggccgg 900
 ccgcgggctt cgtggcgggg cttgggaagg tgcggcccta ctccgtctac ctcttcagct 960
 tctccatgtt cttcaacggc ctgcgggacc tggcgggctc tacggcgggc gactacggcg 1020
 gcctcgtggt cttctgcata ttctttggca tctcctacgg catggtgggg gccctgcagt 1080
 tccaggtgct catggccatc gtgggcaccc acaagtcttc cagtgccatt ggctggtgc 1140
 tgctgatgga ggcggtggcc gtgctcgtcg ggcccccttc gggaggcaaa ctcttgatg 1200
 cgaccacagt ctacatgtac gtgttcattc tggcgggggc cgaggtgctc acctcctccc 1260
 tgattttgct gctgggcaac ttcttctgca ttaggaagaa gcccaaagag ccacagcctg 1320
 aggtggcggc cgcggaggag gagaagctcc acaagcctcc tgcagactcg ggggtggact 1380
 tgcgggaggt ggagcatttc ctgaaggctg agcctgagaa aaacggggag gtggttcaca 1440
 ccccggaac aagtgtctga gtggctgggc ggggcccggc ggcacaggga ggaggtacag 1500
 aagccggcaa cgttgctat ttattttaca aactggactg gctcaggcag ggccacggct 1560
 gggctccagc tgccggccca gcggatcgtc gccgatcag tgttttgagg gggaaggtgg 1620
 cggggtggga accgtgtcat tccagagtgg atctgcgtg aagccaagcc gcaaggttac 1680
 aaggcatcct caccaggggc ccgcctgct gctcccaggt ggctgcggc cactgctatg 1740
 ctcaaggacc tggaaaccca tgcttcgaga caacgtgact ttaatgggag ggtgggtggg 1800
 ccgcagacag gctggcaggg caggtgctgc gtggggccct ctccagcccg tctaccctg 1860
 ggctcacatg gggcctgtgc ccacccctct tgagtgtctt ggggacagct ctttccaccc 1920
 ctggaagatg gaaataaac tgcgtgtggg tggagtgttc tcgtgccgaa ttcaaaaagc 1980
 tt 1982

<210> 140
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004207

<400> 140
cctcttgagt gtcttgggga cagctctttc cacccttgga agatggaaat aaacctgcgt 60

<210> 141
<211> 2054
<212> DNA
<213> Homo sapiens

<300>
<308> NM_004209

<400> 141
cgggaggcgg cagcggctgc agcgttggtg gcatcagcat cagcatcagc ggcagcggca 60
gcggcctcgg gcggggccgg ccggacggac aggcggacag aaggcgccag gggcgcgcg 120
cccggccggg ccggccatgg agggcgctc ctccggcgcg ggcggcgag gggcgccct 180
ggaccccggt agctttgcgc ggcggccca gacctgctc cgggtcgcgt cctgggtgtt 240
ctccatcgcc gtcttcgggc ccatcgtaa cgagggtac gtgaacaccg acagcggccc 300
cgagctgogc tgcgtgttca acgggaacgc gggcgccctc cgcttcggcg tcgcgtggg 360
cctcggagcc ttccctgcct gcgcgcctt cctgctgctc gatgtgcgt tccagcaaat 420
cagcagcgtc cgcgaccgcc ggcgcgcgt gttgctggac ctgggcttct caggactctg 480
gtccttcctg tggttcgtgg gcttctgctt cctcaccaat cagtggcagc gcacggcgcc 540
agggccggcc acgacgcagg cgggggacgc ggcggggcc gccatcgct tcagcttctt 600
ctccatcctc agctgggtgg cgtcacctg gaaggccctg cagcggttcc gcctgggac 660
cgacatgtca ctcttcgcca ccgaacagct gagcaccgg gcgagccagg cctaccccg 720
ctatccggtg ggcagcggcg tggaggcac cgagacctac cagagcccg ccttcaccga 780
gacctggac accagcccca aagggtacca ggtgcccgc tactagcggc tggcaggcac 840
agaccagggc tccaaggcca cccaccaac gcaggccca gggctctcgg gacctccctt 900
gggtccttcc agctcagtgc cgcggacaga gtaggtggcc gctttgcgc atccggggcc 960
aagagggggg ggaccgcgt gtctgggctg cccctgcaa gttccccag tccctcagca 1020
cctggcccca ggaactgagg cctgagaagg ggatagact gccaggacg tgtgtcccta 1080
gcctggaatg gactggcctg gggaggctt tccctcttg ggccacacct gctcactctg 1140
gggttggggg tccagctgcc ctctacgatc aggtgcagg gctgcccagg acaaagcggg 1200
ggcaggggaa agacaccacc ctgcgccaa gactggggat cctggccact gttcccatcc 1260
catgtccctg tgggtagtga ctgtctcgtt tctgtcatg tgggtgcgtc cgtccggagc 1320
cactctccac tttctctcac aggtctctag aacagcccag ccctgtcagt gttgtgatca 1380
tgggtcagtc ttccgggttt acctcctagt actccacaag ctgctcctct ctctgtggcc 1440
ccggccctg ccaggtgtg ggtggttctg gccaggaagg cacaaggtag ctgtgggcca 1500
agacaccagg cctgtcctag cccttcagta agaccttgc aggagaggag aaggatgcct 1560
gggtgccagg caagacaagc ccctcagcag gagagaggcc cagaggctcc agctggccac 1620
cgtgcccac aagatggccc ctgtgtggtt ccctttacct tggcttctg gccagtcct 1680
tgctctcca cctgcacct gcttctggc ccagtcacc gttggagtcc ctctgcatag 1740
ctgactactc atgcattgct caaagctggc ttttcacatt aagtcaacac caaacgtggt 1800
tgccacattt catcagacag acacctccct ctggagatgc agttgagtga caaccttgtt 1860
acattgtagc ctgaccaat tctgtgtgga tatttaagt aacatgttta caatttttgt 1920
atatacact ctctccctct cctgaaagac cagagattgt gtattttcag tgtcccatgt 1980
tccgactgca ccttctttac aataaagact gtaactgagc tgactgtgaa aaaaaaaaa 2040
aaaaaaaaa aaaa 2054

<210> 142
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_004209

<400> 142
gatgcagttg agtgacaacc ttgttacatt gtagcctaga ccaattctgt gtggatattt 60

<210> 143
 <211> 1224
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004217

<400> 143
 ggccgggaga gtagcagtgc cttggacccc agctctcctc cccctttctc tctaaggatg 60
 gcccagaagg agaactccta cccctggccc tacggccgac agacggctcc atctggcctg 120
 agcaccctgc cccagcgagt cctccggaaa gagcctgtca ccccatctgc acttgtcctc 180
 atgagccgct ccaatgtcca gccacagct gccctggcc agaaggtgat ggagaatagc 240
 agtgggacac ccgacatctt aacgcggcac ttcacaattg atgactttga gattgggcgt 300
 cctctgggca aaggcaagtt tggaaacgtg tacttggctc gggagaagaa aagccatttc 360
 atcgtggcgc tcaaggtcct cttcaagtcc cagatagaga aggagggcgt ggagcatcag 420
 ctgcgcagag agatcgaaat ccaggccac ctgcaccatc ccaacatcct gcgtctctac 480
 aactattttt atgaccggag gaggatctac ttgattctag agtatgcccc ccgcggggag 540
 ctctacaagg agctgcagaa gagctgcaca tttgacgagc agcgaacagc cacgatcatg 600
 gaggagttgg cagatgctct aatgtactgc catgggaaga aggtgattca cagagacata 660
 aagccagaaa atctgctctt agggctcaag ggagagctga agattgctga cttcggctgg 720
 tctgtgcatg cgccctccct gaggaggaag acaatgtgtg gcaccctgga ctacctgcc 780
 ccagagatga ttgagggcg catgcacaat gagaaggtgg atctgtggtg cattggagtg 840
 ctttgctatg agctgtcgtt ggggaaccca ccttttgaga gtgcatcaca caacgagacc 900
 tatcgccgca tcgtcaaggt ggacctaaag ttccccgctt ctgtgcccac gggagcccag 960
 gacctcatct ccaaactgct caggcataac ccctcggaac ggctgcccct ggcccaggct 1020
 tcagcccacc cttgggtccg ggccaactct cggaggggtg tgccctccctc tgcccttcaa 1080
 tctgtgcct gatgtccct gtcattcact cgggtgctg tggtttgtatg tctgtgtatg 1140
 tataggggaa agaaggatc cctaactgtt cccttatctg ttttctacct cctcctttgt 1200
 ttaataaagg ctgaagcttt ttgt 1224

<210> 144
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004217

<400> 144
 gtctgtgtat gtatagggga aagaaggat ccctaactgt tcccttatct gttttctacc 60

<210> 145
 <211> 983
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004335

<400> 145
 gtggaattca tggcatctac ttctgtatgac tattgcagag tgcccatgga agacggggat 60
 aagcgtgtga agcttctgct ggggatagga attctgggtc tcctgatcat cgtgattctg 120
 ggggtgccct tgattatctt caccatcaag gccaacagcg aggcctgccg ggacggcctt 180
 cgggcagtga tggagtgtcg caatgtcacc catctcctgc aacaagagct gaccgaggcc 240
 cagaagggtt ttcaggatgt ggaggcccag gccgccacct gcaaccacac tgtgatggcc 300
 ctaatggctt ccctggatgc agagaaggcc caaggacaaa agaaagtgga ggagcttgag 360
 ggagagatca ctacattaaa ccataagctt caggacgcgt ctgcagaggt ggagcgactg 420
 agaagagaaa accaggtctt aagcgtgaga atcgcggaaga agaagtacta ccccgactcc 480
 caggactcca gctccgctgc ggcgcccag ctgctgattg tgctgctggg cctcagcgct 540
 ctgctgcagt gagatcccag gaagctggca catcttgaa ggtccgtcct gctcggcttt 600

```

tgcgttgaac attcccttga tctcatcagt tctgagcggg tcatggggca acacgggttag 660
cggggagagc acggggtagc cggagaaggg cctctggagc aggtctggag gggccatggg 720
gcagtcctgg gtgtggggac acagtcgggt tgacccaggg ctgtctccct ccagagcctc 780
cctccggaca atgagtcctc cctcttgtct cccaccctga gattgggcat ggggtgcggg 840
gtggggggca tgtgctgcct gttgttatgg gttttttttg cggggggggg tgcttttttc 900
tgggggtcttt gagctccaaa aaataaacac ttcctttgag ggagagcaaa aaaaaaaaaa 960
aaaaaaaaaa aaaaaaaaaa aaa 983

```

<210> 146

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004335

<400> 146

```

ggttgctttt ttctggggtc tttgagctcc aaaaaataaa cacttccttt gagggagagc 60

```

<210> 147

<211> 3446

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004336

<400> 147

```

ttctagtttg cggttcaggt ttgccgtgc cggccagcgt cctctggcca tggacacccc 60
ggaaaatgtc cttcagatgc ttgaagccca catgcagagc tacaagggca atgaccctct 120
tggtgaatgg gaaagataga tacagtgggt agaagagaat ttctctgaga ataaagaata 180
cttgataact ttactagaac atttaatgaa ggaattttta gataagaaga aataccacaa 240
tgacccaaga ttcacaggtt attgtttaaa atttgctgag tacaacagtg acctccatca 300
atTTTTtTgag ttctgttaca accatgggat tggaaacctg tcatccctc tgtacattgc 360
ctgggcgggg catctggaag cccaaggaga gctgcagcat gccagtgtg tctctcagag 420
aggaattcaa aaccaggctg aaccagagga gttcctgcaa caacaataca ggttatttca 480
gacacgcctc actgaaaccc atttgccagc tcaagctaga acctcagaac ctctgcataa 540
tgttcaggtt ttaaatacaa tgataacatc aaatcaaat ccaggaaata acatggcctg 600
catttctaag aatcagggtt cagagcttct tggagtgtata tcttcagctt gtgataaaga 660
gtcaaatatg gaacgaagag tgatcacgat ttctaaatca gaatattctg tgcactcatc 720
tttggcatcc aaagtgtgat ttgagcaggt tggtatgtat tgcaaggaga agcttattcg 780
tggggaatca gaattttcct ttgaagaatt gagagcccag aaatacaatc aacggagaaa 840
gcatgagcaa tgggtaaaatg aagacagaca ttatatgaaa aggaaagaag caaatgcttt 900
tgaagaacag ctattaaaac agaaaatgga tgaacttcat aagaagttgc atcagggtgt 960
ggagacatcc catgaggatc tgcccgcttc ccaggaaagg tccgaggtta atccagcacg 1020
tatggggcca agtgtaggct cccagcagga actgagagcg ccattgtctc cagtaacctc 1080
tcagcagaca ccagtgaaca tggaaaagaa cccaagagag gcacctcctg ttgttcctcc 1140
tttgccaaat gctattttctg cagcttttgt gtccccagcc accagccaga gcattgctcc 1200
tctgtttcct ttgaaagccc agacagtaac agactccatg tttgcagtgg ccagcaaaga 1260
tgctggatgt gtgaataaga gtactcatga attcaagcca cagagtggag cagagatcaa 1320
agaaggggtgt gaaacacata aggttgccaa cacaagttct ttccacacaa ctccaaacac 1380
atcactggga atggttcagg caacgccatc caaagtgcag ccatacccca ccgtgcacac 1440
aaaaagaagca ttagttttca tcatgaatat gtttcaggct cctacacttc ctgatatttc 1500
tgatgacaaa gatgaatggc aatctctaga tcaaaatgaa gatgcatttg aagcccagtt 1560
tcaaaaaaat gtaaggtcat ctggggcttg gggagtcaat aagatcatct cttctttgtc 1620
atctgctttt catgtgtttg aagatggaaa caaagaaaat tatggattac cacagcctaa 1680
aaataaaacc acaggagcca ggaccttgga agaacgctct gtcagcagac ttccttcaaa 1740
accaaaaggag gaagtgcctc atgctgaaga gtttttggat gactcaactg tatggggtat 1800
tcgtgtcaac aaaaccctgg caccagctcc taagagccca ggagacttca catctgctgc 1860
acaacttgcg tctacacatc tccacaagct tccagtggag tcaagtgcaca ttttagaaga 1920
taaaagaaat gtggtagcaa aacagtgtag ccaggcgact ttggattctt gtgaggaaaa 1980

```



```

catgggtgggtg ccttcaaggg atggaaaatt cagtccaatt caagagaaaa gcccaaaaaca 2040
ggccttggtcg tctcacatgt attcagcatc cttacttcgt ctgagccagc ctgctgcagg 2100
tgggggtactt acctgtgagg cagagttggg cgttgaggct tgcagactca cagacactga 2160
cgctgccatt gcagaagatc caccagatgc tattgtctgg ctccaagcag aatggatgca 2220
gatgagttca cttgggactg ttgatgctcc aaacttcatt gttgggaacc catgggatga 2280
taagctgatt ttcaaacttt tatctgggct ttctaaacca gtgagttcct atccaaatac 2340
ttttgaatgg caatgtaaac ttccagccat caagcccaag actgaatttc aattgggttc 2400
taagctggtc tatgtccatc accttcttgg agaaggagcc tttgcccagg tgtacgaagc 2460
taccagggga gatctgaatg atgctaaaaa taaacagaaa tttgttttaa aggtccaaaa 2520
gcctgccaac ccctgggaat tctacattgg gaccagttg atggaaagac taaagccatc 2580
tatgcagcac atgtttatga agttctatc tgcccactta ttccagaatg gcagtgtatt 2640
agtaggagag ctctacagct atggaacatt attaaatgcc attaacctct ataaaaatac 2700
ccctgaaaaa gtgatgcctc aaggtcttgt catctctttt gctatgagaa tgctttacat 2760
gattgagcaa gtgcatgact gtgaaatcat tcatggagac attaaaccag acaatttcac 2820
acttggaaac ggatttttgg aacaggatga tgaagatgat ttatctgctg gcttggcact 2880
gattgacctg ggtcagagta tagatatgaa actttttcca aaaggaacta tattcacagc 2940
aaagtgtgaa acatctggtt ttcagtgtgt tgagatgctc agcaacaaac catggaacta 3000
ccagatcgat tactttgggg ttgctgcaac agtatattgc atgctctttg gcacttacct 3060
gaaagtgaag aatgaaggag gagagtgtaa gcctgaaggt ctttttagaa ggcttcctca 3120
tttgatatg tggaatgaat ttttcatgt tatgttgaat attccagatt gtcacatct 3180
tccatctttg gattttgttaa ggcaaaagct gaagaaagta tttcaacaac actatactaa 3240
caagattagg gccctacgta ataggcta at tgtactgctc ttagaatgta agcgttcacg 3300
aaaataaaat ttgatatatg acagtcctta aaaatcacac tgtaaatatg aatctgctca 3360
ctttaaacct gttttttttt catttattgt ttatgtaaat gtttgttaaa aataaatccc 3420
atggaatatt tccatgtaaa aaaaaa 3446

```

<210> 148

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004336

<400> 148

```

ttagggccct acgtaatagg ctaattgtac tgctcttaga atgtaagcgt tcacgaaaat 60

```

<210> 149

<211> 739

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004345

<400> 149

```

taaagcaaac ccagccac accctggcag gcagccagg atgggtggat caggaaggct 60
cctggttggg cttttgcac aggtcaggc tgggcataaa ggaggctcct gtgggctaga 120
gggaggcaga catggggacc atgaagacc aaagggatgg ccaactccctg gggcggtggt 180
cactggtgct cctgctgctg ggcctggtga tgccctctgg catcattgcc caggctcctca 240
gtacaagga agctgtgctt cgtgctatag atggcatcaa ccagcgggtcc tcggatgcta 300
acctctaccg cctcctggac ctggaccca ggcccacgat ggatggggac ccagacacgc 360
caaagcctgt gagcttcaca gtgaaggaga cagtgtgccc caggacgaca cagcagtcac 420
cagaggattg tgacttcaag aaggacgggc tgggtgaagc gtgtatgggg acagtgaccc 480
tcaaccaggc caggggctcc tttgacatca gttgtgataa ggataacaag agatttgccc 540
tgctgggtga tttcttccgg aaatctaaag agaagattgg caaagagttt aaaagaattg 600
tccagagaat caaggatttt ttgcggaatc ttgtaccag gacagagtc tagtggtgct 660
cctaccctgg ctgaggttc tgggctctga gaaataaact atgagagcaa tttcaaaaaa 720
aaaaaaaaa aaaaaaaaaa 739

```

<210> 150
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_004345

<400> 150
gcaaagagtt taaaagaatt gtccagagaa tcaaggattt tttgcggaat cttgtaccca 60

<210> 151
<211> 1432
<212> DNA
<213> Homo sapiens

<300>
<308> NM_004577

<400> 151
gaggaaaatt cttccagcga tgggtctccca ctcagagctg aggaagcttt tctactcagc 60
agatgctgtg tgttttgatg ttgacagcac ggatcatcaga gaagaaggaa tcgatgagct 120
agccaaaatc tgtggcgttg aggacgcggt gtcagaaatg acacggcgag ccatgggagg 180
ggcagtgccct ttcaaagctg ctctcacaga gcgcttagcc ctcatccagc cctccaggga 240
gcaggtgcag agactcatag cagagcaacc cccacacctg acccccggca taaggagct 300
ggtaagtgcg ctacaggagc gaaatgttca ggttttccta atatctggtg gctttaggag 360
tattgtagag catgttgctt caaagctcaa tatcccagca accaatgtat ttgccaatag 420
gctgaaatc tactttaacg gtgaatatgc aggttttgat gagacgcagc caacagctga 480
atctggtgga aaaggaaaag tgattaaact tttaaaggaa aaatttcatt ttaagaaaat 540
aatcatgatt ggagatggtg ccacagatat ggaagcctgt cctcctgctg atgctttcat 600
tggatttgga ggaaatgtga tcaggcaaca agtcaaggat aacgccaat ggtatatcac 660
tgattttgta gagctgctgg gagaactgga agaataacat ccattgtcgt acagctocaa 720
acaacttcag atgaattttt acaagttata cagattgata ctgtttgctt acagttgcct 780
attacaactt gctatagaaa gttggtacaa atgatctgta ctttaacta cagttaggaa 840
tcctagaaga ttgctttttt ttttttttta actgtagtcc cagtattata tgatgactat 900
tgatttcctg gagaggtttt tttttttttt gagacagaat cttgctctgt tgcccaggct 960
ggagtcgagt ggcgcggtct cggctcactg caagctctgc ctcccagggt cagccattc 1020
tcctgcctca gcctcccag tagctgggac tacaggcacc cgccaccaca tccggcta 1080
tttttgatatt tttagtagag acgggggttg acogtggttag ccaggatggt cttgatctcc 1140
tgacctgtg atccgcctgc ctacagcctcc caaagtgtg ggattacagg cttggggccac 1200
cgcgcccagc caatgtccta gagagttttg tgatctgaat tctttatgta tattttagc 1260
tatatttcac acaaagtgt ttaagtgtg agagtcaatt aaacacctt actcttagaa 1320
atacggatc ggcagcctc agtgaatatt ggtttctctt tggatgtca ataaaagttt 1380
atccgtatgt cagaacggat ttgtggaaaa aaaaaaaaaa aaaaaaaaaa aa 1432

<210> 152
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_004577

<400> 152
tagaaatcag gattcggcag ccttcagtga atattggttt ctctttggta tgtcaataaa 60

<210> 153
<211> 1530
<212> DNA
<213> Homo sapiens

<300>

<308> NM_004701

<400> 153

```

aatcctggaa caaggtaca gcgtcgaaga tccccagcgc tgcgggctcg gagagcagtc 60
ctaacggcgc ctctgacgt agtgtcctcc cttttcagtc cgcgtccctc cctgggcccgg 120
gctggcactc ttgccttccc cgtccctcat ggcgctgctc cgacgcccga cgggtgtccag 180
tgatttggag aatattgaca caggagttaa ttctaaagtt aagagtcagtg tgactattag 240
gcgaactggt ttagaagaaa ttggaaatag agttacaacc agagcagcac aagtagctaa 300
gaaagctcag aacaccaaag ttccagttca acccaccaaa acaacaaatg tcaacaaaca 360
actgaaacct actgcttctg tcaaaccagt acagatggaa aagttggctc caaagggtcc 420
ttctccaca cctgaggatg tctccatgaa ggaagagaat ctctgccaa ctttttctga 480
tgcttctgctc tgcaaaatcg aggacattga taacgaagat tgggagaacc ctacagctctg 540
cagtgcactac gttaaggata tctatcagta tctcaggcag ctggagggtt tgcagtcctat 600
aaacccacat ttcttagatg gaagagatat aaatggacgc atgcgtgcc a tcttagtgga 660
ttggctggta caagtccact ccaagtttag gcttctgcag gagactctgt acatgtgcgt 720
tggcattatg gatcgatttt tacaggttca gccagtttcc cggaagaagc ttcaattagt 780
tggttact gctctgctct tggcttccaa gtatgaggag atgttttctc caaatattga 840
agactttgtt tacatcacag acaatgctta taccagttcc caaatccgag aaatggaaac 900
tctaattttg aaagaattga aatttgagtt gggctgaccc ttgccactac acttcttaag 960
gcgagcatca aaagccgggg aggttgatgt tgaacagcac actttagcca agtatttgat 1020
ggagctgact ctcatcgact atgatatggt gcattatcat ccttctaagg tagcagcagc 1080
tgcttctgctc ttgtctcaga aggttctagg acaaggaaaa tggaaactta agcagcagta 1140
ttacacagga tacacagaga atgaagtatt ggaagtcag cagcacatgg ccaagaatgt 1200
ggtgaaagta aatgaaaact taactaaatt catgccatc aagaataagt atgcaagcag 1260
caaactcctg aagatcagca tgatccctca gctgaaactca aaagccgtca aagacctgct 1320
ctcccaactg ataggaagggt cctaggctgc cgtgggccct ggggatgtgt gcttcattgt 1380
gccctttttc ttattggttt agaactcttg attttgtaca tagtcctctg gtctatctca 1440
tgaaacctct tctcagacca gttttctaaa catatatgta ggaaaaataa agcgattggt 1500
ttttcttaag gtaaaaaaaa aaaaaaaaaa 1530

```

<210> 154

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004701

<400> 154

```

agaactcttg attttgtaca tagtcctctg gtctatctca tgaaacctct tctcagacca 60

```

<210> 155

<211> 2536

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004702

<400> 155

```

agcgggtgcg gggcgggacc ggcccggcct atatattggg ttggcgccgg cgccagctga 60
gccgagcggg agctggctcg gcgaggtttt atacacctga aagaagagaa tgtcaagacg 120
aagtagccgt ttacaagcta agcagcagcc ccagcccagc cagacggaat cccccaaga 180
agcccagata atccaggcca agaagaggaa aactaccag gatgtcaaaa gaagtctggc 240
taaactatgt aaaaaaggag agcagatatg ttcatgacaa acattttgaa gttctgcatt 300
ctgacttgga accacagatg aggtccatac ttctagactg gcttttagag gtatgtgaag 360
tatacacact tcatagggaa acattttatc ttgcacaaga cttttttgat agatttatgt 420
tgacacaaa ggatataaat aaaaatatgc ttcaactcat tggaaattacc tcattattca 480
ttgcttccaa acctgaggaa atctatgctc ctaaactcca agagtttgct tacgtcactg 540
atggtgcttg cagtgaagag gatattctaa ggatggaact cattatatta aaggctttaa 600

```

```

aatgggaact ttgtcctgta acaatcatct cctgggctaaa tctctttctc caagttgatg 660
ctcttaaaaga tgctcctaaa gttcttctac ctcagtattc tcaggaaaca ttcattcaaa 720
tagctcagct tttagatctg tgtattctag ccattgattc attagagttc cagtacagaa 780
tactgactgc tgctgccttg tgccatttta cctccattga agtggttaag aaagcctcag 840
gtttggagtg ggacagtatt tcagaatgtg tagattggat ggtacctttt gtcaatgtag 900
taaaaagtac tagtccagtg aagctgaaga cttttaagaa gattcctatg gaagacagac 960
ataatatcca gacacataca aactattttg ctatgctgga ggaagtaaat tacataaaca 1020
ccttcagaaa aggggggacag ttgtcaccag tgtgcaatgg aggcattatg acaccaccga 1080
agagcactga aaaaccacca ggaaaacact aaagaagata actaagcaaa caagttggaa 1140
ttcaccaaga ttgggtagaa ctggtatcac tgaactacta aagttttaca gaaagtagtg 1200
ctgtgattga ttgccctagc caattcacaa gttacactgc cattctgatt ttaaaactta 1260
caattggcac taaagaatac atttaattat ttccatgttt agctgttaaa gaaacagcag 1320
gacttggtta caaagatgtc ttcattccca aggttactgg atagaagcca accacagtct 1380
ataccatagc aatgtttttc ctttaatcca gtgttactgt gtttatcttg ataaactagg 1440
aattttgtca ctggagtttt ggactggata agtgctacct taaagggat actaagtgat 1500
acagtacttt gaactagtt gttagattct caaaattcct acactcttga ctagtgcaat 1560
ttggttcttg aaaattaaat ttaaacctgt ttacaaaggt ttagttttgt aataaggtga 1620
ctaatttctc tatagctgct atagcaagct attataaaac ttgaatttct acaaaggtg 1680
aaatttaagt ttttttaaac tagtttattt gccttgccat aacacatttt ttaactaata 1740
aggcttagat gaacatggtg ttcaacctgt gctctaaaca gtgggagtag caaagaaatt 1800
ataaacaaga taaatgctgt ggctccttcc taactggggc tttcttgaca tgtaggttgc 1860
ttgtaataaa cttttttgta tatcacaatt tgggtgaaaa acttaagtag cctttcaaac 1920
tatttatatg aggaagtcac tttactactc taagatatcc ctaaggaatt ttttttttta 1980
atttagtgtg actaaggctt tatttatgtt tgtgaaactg ttaaggctct ttctaaattc 2040
ctccattgtg agataaggac agtgtcaaa tgataaagct taacacttga cctaaacttc 2100
tattttctta aggaagaaga gtattaaata tatactgact cctagaaatc tatttattaa 2160
aaaaagacat gaaaacttgc tgtacatagg ctagctatct ctaaatattt taaattagct 2220
tttctaaaaa aaaaatccag cctcataaag tagattagaa aactagattg ctagtattt 2280
ttgttatcag atatgtgaat ctcttctccc tttgaagaaa ctatacattt attgttacgg 2340
tatgaagtct tctgtatagt ttgtttttta actaatattt gtttcagtat ttgtctgaa 2400
aagaaaaccac cactaattgt gtacatagtg attatataaa cttaaccttt taatactgtt 2460
tatttttagc ccattgttta aaaaataaaa gttaaaaaaa ttaactgct taaaagtaaa 2520
aaaaaaaaa aaaaaa 2536

```

<210> 156

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004702

<400> 156

```

gtttgtgaaa ctgttaaggt cctttctaaa ttccctcatt gtgagataag gacagtgtca 60

```

<210> 157

<211> 1491

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004710

<400> 157

```

gcggcgccgg cagcgccggc gacggcgaca tggagagcgg ggcctacggc gcggccaagg 60
cgggcgccgc cttcgacctg cggcgcttcc tgacgcagcc gcagggtggt gcgcgcgccg 120
tgtgcttggt cttcgccctg atcgtgttct cctgcactca tggtgagggc tacagcaatg 180
cccacgagtc taagcagatg tactgcgtgt tcaaccgcaa cgaggatgcc tgccgctatg 240
gcagtgccat cgggggtgctg gccttctctg cctcggcctt cttcttggtg gtcgcgcggt 300
atctcccca gatcgaac gccactgacc gcaagtacct ggtcattggt gacctgctct 360
tctcagctct ctggaccttc ctgtgggttg ttgggttctg cttcctcacc aaccagtggtg 420

```

```

cagtcaccaa cccgaaggac gtgctggtgg gggccgactc tgtgagggca gccatcacct 480
tcagcttctt ttccatcttc tcctgggggtg tgctggcctc cctggcctac cagcgctaca 540
aggctggcgt ggacgacttc atccagaatt acgttgacct cactccggac cccaacactg 600
cctacgcctc ctaccaggt gcatctgtgg acaactacca acagccacct ttcaccaga 660
acgcggagac caccgagggc taccagccgc ccctgtgtgta ctgagcggcg gttagcgtgg 720
gaagggggac agagagggcc ctcccctctg ccctggactt tcccatgagc ctccctggaac 780
tgccagcccc tctctttcac ctgttccatc ctgtgcagct gacacacagc taaggagcct 840
catagcctgg cgggggctgg cagagccaca cccaagtgc ctgtgccagc agggcttcag 900
tcagcgctc actcctccag ggcattttta ggaaagggtt ttcagctagt gtttttctc 960
gcttttaatg acctcagccc cgctgcagc ggctagaagc cagcaggtgc ccatgtgcta 1020
ctgacaagtg cctcagcttc cccccggccc gggtcaggcc gtgggagccg ctattatctg 1080
cgttctctgc caaagactcg tgggggccat cacacctgcc ctgtgcagcg gagccggacc 1140
aggctcttgt gtcctcactc aggtttgctt ccctgtgcc cactgctgta tgatctgggg 1200
gccaccaccc tgtgccggtg gcctctgggc tgccctcccg ggtgtgaggg cggggctggt 1260
gctcatggca ctctctcctt gctcccaccc ctggcagcag ggaagggtt tgcccgacaa 1320
caccagctt tatgtaaata ttctgcagtt gttacttagg aagcctgggg agggcagggg 1380
tgcccatgg ctcccagact ctgtctgtgc cgagtgtatt ataaaatcgt gggggagatg 1440
cccgcctgg gatgctgttt ggagacggaa taaatgtttt ctcattcagt a 1491

```

<210> 158

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004710

<400> 158

```

ttgcctgaca acaccagct ttatgtaaat attctgcagt tggtacttag gaagcctggg 60

```

<210> 159

<211> 3324

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004856

<400> 159

```

gcagagcacc gcgccttagc cgcgaagttc tagttcttgc tgccggctcct aacgtcccgc 60
agtcttgcgc agccagccgt cccgcagcgc cgtttgggcg gcgtggagcc tgctgccatg 120
aagtccagca gagctaagac accccggaac cctaccgtga aaaaagggtc ccaaacgaac 180
cttaagagac cagttggggt atactgtagg gtgcgcccac tgggctttcc tgatcaagag 240
tggtgcatag aagtgatcaa taatacaact gttcagcttc atactcctga gggctacaga 300
ctcaaccgaa atgggagact taaggagact cagtattcat ttaaacaagt atttggcact 360
cacaccaccc agaaggaact ctttgatgtt gtggctaata ccttgggtcaa tgacctcatt 420
catggcaaaa atggctcttc ttttacatat ggtgtgacgg gaagtggaaa aactcacaca 480
atgactgggt ctccagggga aggagggctg ctccctcggt gtttggacat gatctttaac 540
agtatagggt catttcaagc taaacgatat gttttcaaat ctaatgatag gaatagtatg 600
gatatacagt gtgaggttga tgccttatta gaacgtcaga aaagagaagc tatgcccaat 660
ccaaagactt cttctagcaa acgacaagta gatccagagt ttgcagatat gataactgta 720
caagaattct gcaaagcaga agaggttgat gaagatagtg tctatggtgt atttgtctct 780
tatattgaaa tatataataa ttacatatat gatctattgg aagaggtgcc gtttgatccc 840
ataaaacca aacctccaca atctaaattg cttcgtgaag ataagaacca taacatgtat 900
gttgaggat gtacagaagt tgaagtgaat tctactgagg aggcctttga agttttctgg 960
agaggccaga aaaagagacg tattgctaata acccatttga atcgtgagtc cagccgttcc 1020
catagcgtgt tcaacattaa attagttcag gctcccttgg atgcagatgg agacaatgtc 1080
ttacaggaaa aagaacaaat cactataagt cagttgtcct tggtagatct tgctggaagt 1140
gaaagaacta accggaccag agcagaaggg aacagattac gtgaagctgg taatattaat 1200
cagtcactaa tgacgctaag aacatgtatg gatgtcctaa gagagaacca aatgtatgga 1260
actaacaaga tggttccata tcgagattca aagttaacc atctgttcaa gaactacttt 1320

```

gatggggaag	gaaaagtgcg	gatgatcgtg	tgtgtgaacc	ccaaggctga	agattatgaa	1380
gaaaacttgc	aagtcattgag	atttgcggaa	gtgactcaag	aagttgaagt	agcaagacct	1440
gtagacaagg	caatatgtgg	tttaacgcct	gggaggagat	acagaaacca	gcctcgaggt	1500
ccagttggaa	atgaaccatt	ggttactgac	gtggttttgc	agagttttcc	acctttgccg	1560
tcatgcgaaa	ttttggatat	caacgatgag	cagacacttc	caaggctgat	tgaagcctta	1620
gagaacgac	ataacttacg	acaaatgatg	attgatgagt	ttaacaaaca	atctaattgct	1680
tttaaaagctt	tggtacaaga	atttgacaat	gctgttttaa	gtaaagaaaa	ccacatgcaa	1740
gggaaactaa	atgaaaagga	gaagatgatc	tcaggacaga	aattggaaat	agaacgactg	1800
gaaaagaaaa	acaaaacttt	agaatataag	attgagattt	tagagaaaac	aactactatc	1860
tatgaggaag	ataaacgcaa	tttgcaacag	gaacttgaaa	ctcagaacca	gaaacttcag	1920
cgacagtttt	ctgacaaacg	cagatttagaa	gccagggttc	aaggcatggt	gacagaaacg	1980
acaatgaagt	gggagaaaga	atgtgagcgt	agagtggcag	ccaaacagct	ggagatgcag	2040
aataaactct	gggttaaaga	tgaaaagctg	aaacaactga	aggctattgt	tactgaacct	2100
aaaactgaga	agccagagag	accctctcgg	gagcgagatc	gagaaaaagt	tactcaaaga	2160
tctgtttctc	catcacctgt	gcctttactc	tttcaacctg	atcagaacgc	accaccaatt	2220
cgtctccgac	acagacgatc	acgctctgca	ggagacagat	gggtagatca	taagcccgcc	2280
tctaaccatgc	aaactgaaac	agtcattgcag	ccacatgtcc	ctcatgccat	cacagtatct	2340
gttgcaaatg	aaaaggcact	agctaagtgt	gagaagtaca	tgctgaccca	ccaggaacta	2400
gcctccgatg	gggagattga	aactaaacta	attaagggtg	atatttataa	aacaaggggt	2460
ggtggacaat	ctgttcagtt	tactgatatt	gagactttta	agcaagaatc	accaaattgt	2520
agtcgaaaaac	gaagatcttc	cacagtagca	cctgcccaac	cagatggtgc	agagtctgaa	2580
tggaccgatg	tagaaacaag	gtgttctgtg	gctgtggaga	tgagagcagg	atcccagctg	2640
ggacctggat	atcagcatca	cgcacaaccc	aagcgcaaaa	agccatgaac	tgacagtccc	2700
agtaactgaaa	gaacattttc	atttgtgtgg	atgatttctc	gaaagccatg	ccagaagcag	2760
tcttccaggt	catcttgtag	aactccagct	ttgttgaaaa	tcacggacct	cagctacatc	2820
atacactgac	ccagagcaaa	gctttcccta	tggttccaaa	gacaactagt	attcaacaaa	2880
ccttgtatag	tatatgtttt	gccatattta	atattaatag	cagaggaaga	ctcctttttt	2940
catcactgta	tgaatttttt	ataatgtttt	tttaaaatat	atttcatgta	tacttataaa	3000
ctaattcaca	caagtgtttg	totttagatga	ttaaggaaga	ctatatctag	atcatgtctg	3060
attttttatt	gtgacttctc	cagccctggg	ctgaatttct	taagggtttta	taaacaaatg	3120
ctgtctattta	ttagctgcaa	gaatgcactt	tagaactatt	tgacaattca	gactttcaaa	3180
ataaagatgt	aatgactggg	ccaataataa	ccatttttagg	aagggtgtttt	gaattctgta	3240
tgtatatatt	cactttctga	catttagata	tgccaaaaga	attaaaatca	aaagcactaa	3300
gaaataaaaa	aaaaaaaaaa	aaaa	3324			

<210> 160

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004856

<400> 160

caaagctttc	cctatggttc	aaagacaact	agtattcaac	aaaccttgta	tagtgtatgt	60
------------	------------	------------	------------	------------	------------	----

<210> 161

<211> 1536

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004900

<400> 161

acagagcttc	aaaaaaagag	cgggacaggg	acaagcgtat	ctaagaggct	gaacatgaat	60
ccacagatca	gaaatccgat	ggagcggatg	tatcgagaca	cattctacga	caactttgaa	120
aacgaaccct	tcctctatgg	tcggagctac	acttggctgt	gctatgaagt	gaaaataaag	180
aggggcogct	caaatctcct	ttgggacaca	ggggtctttc	gaggccaggt	gtatttcaag	240
cctcagtacc	acgcagaaat	gtgcttcctc	tcttggttct	gtggcaacca	gctgcctgct	300
tacaagtgtt	tcagatcac	ctggtttgta	tcctggacct	cctgcccga	ctgtgtggcg	360

```

aagctggccg aattcctgtc tgagcacccc aatgtcaccg tgaccatctc tgccgcccgc 420
ctctactact actgggaaag agattaccga agggcgctct gcaggctgag tcaggcagga 480
gcccgcgtga cgatcatgga ctatgaagaa tttgcatact gctgggaaaa ctttgtgtac 540
aatgaaggtc agcaattcat gccttggtac aaattcgatg aaaattatgc attcctgcac 600
cgcacgctaa aggagattct cagatacctg atggatccag acacattcac tttcaacttt 660
aataatgacc ctttggctct tcgacggcgc cagacctact tgtgctatga ggtggagcgc 720
ctggacaatg gcacctgggt cctgatggac cagcacatgg gctttctatg caacgaggct 780
aagaatcttc tctgtggctt ttacggccgc catgcgagc tgcgcttctt ggacctggtt 840
ccttctttgc agttggaccc ggcccagatc tacagggtca cttggttcat ctccctggagc 900
ccctgcttct cctggggctg tgccggggaa gtgctgctg tccctcagga gaacacacac 960
gtgagactgc gcactctcgc tgcccgcatc tatgattacg accccctata taaggaggcg 1020
ctgcaaatgc tgcgggatgc tggggcccaa gtctccatca tgacctacga tgagtttgag 1080
tactgctggg acacctttgt gtaccgccag ggatgtccct tccagccctg ggatggacta 1140
gaggagcaca gccaagccct gagtgggagg ctgcccccca ttctccagaa tcagggaaac 1200
tgaaggatgg gcctcagtct ctaagggaag cagagacctg ggttgagcag cagaataaaa 1260
gatctcttc caagaaatgc aaacagaccg ttcaccacca tctccagctg ctcacagaca 1320
ccagcaaacg aatgtgctcc tgatcaagta gattttttta aaatcagagt caattaattt 1380
taattgaaaa tttctcttat gttccaagtg tacaagagta agattatgct caatattccc 1440
agaatagttt tcaatgtatt aatgaagtga ttaattggct ccatatttag actaataaaa 1500
cattaagaat ctcccataat tgtttccaca aacact 1536

```

<210> 162
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004900

```

<400> 162
tgctcacaga caccagcaaa gcaatgtgct cctgatcaag tagatttttt aaaaatcaga 60

```

<210> 163
 <211> 1722
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_004988

```

<400> 163
cgtagagttc ggccgaagga acctgaccca ggctctgtga ggaggcaagg ttttcagggg 60
acaggccaac ccagaggaca ggattccctg gaggccacag aggagacca aggagaagat 120
ctgcctgtgg gtcttcattg cccagctcct gccacactc ctgcctgctg ccctgacgag 180
agtcatcatg tctcttgagc agaggagtct gcaactgcaag cctgaggaag cccttgaggc 240
ccaacaagag gccctgggcc tgggtgtgtg gcaggctgcc gcctcctcct cctctcctct 300
ggctcctggc accctggagg aggtgccac tgctgggtca acagatcctc cccagagtcc 360
tcagggagcc tccgcctttc ccactaccat caacttcaact cgacagaggc aaccagtgta 420
gggttccagc agccgtgaag aggaggggcc aagcacctct tgtatcctgg agtccttggt 480
ccgagcagta atcactaaga aggtggctga tttgggtggg tttctgctcc tcaaatatcg 540
agccagggag ccagtcacaa aggcagaaat gctggagagt gtcatacaaaa attacaagca 600
ctgttttctc gagatcttcg gcaaagcctc tgagtcttg cagctggtct ttggcattga 660
cgtgaaggaa gcagacccca ccggccactc ctatgtcctt gtcacctgcc taggtctctc 720
ctatgatggc ctgctgggtg ataatcagat catgcccagg acaggcttcc tgataattgt 780
cctggtcatg attgcaatgg agggcggcca tgctcctgag gaggaaatct gggaggagct 840
gagtgatgat gaggtgatg atgggaggga gcacagtgcc tatggggagc ccaggaagct 900
gctcacccaa gatttggtgc aggaaaagta cctggagtac cggcaggtgc cggacagtga 960
tcccgcacgc tatgagttcc tgtgggggtc aagggccctt gctgaaacca gctatgtgaa 1020
agtcttgag tatgtgatca aggtcagtgc aagagttcgc tttttcttcc catccctgcg 1080
tgaagcagct ttgagagag aggaagagg agtctgagca tgagttgcag ccaggggccag 1140
tgggaggggg actggggccag tgcaccttcc agggccgcgt ccagcagctt cccctgcctc 1200

```

```

gtgtgacatg aggcccatte ttcactctga agagagcggg cagtgttctc agtagtaggt 1260
ttctgttcta ttgggtgact tggagattta tctttgttct cttttggaat tgttcaaagt 1320
ttttttttta agggatgggt gaatgaactt cagcatccaa gtttatgaat gacagcagtc 1380
acacagttct gtgtatatag ttttaagggt agagtcctgt gttttattca gattgggaaa 1440
tccattctat tttgtgaatt gggataataa cagcagtggg ataagtactt agaaatgtga 1500
aaaatgagca gtaaaataga tgagataaag aactaaagaa attaaagat agtcaattct 1560
tgccttatac ctcagtctat tctgtaaaat ttttaaagat atatgcatac ctggatttcc 1620
ttggcttctt tgagaatgta agagaaatta aatctgaata aagaattctt cctgttaaaa 1680
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 1722

```

<210> 164

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004988

<400> 164

```

cagattggga aatccattct attttgtgaa ttgggataat aacagcagtg gaataagtac 60

```

<210> 165

<211> 2334

<212> DNA

<213> Homo sapiens

<300>

<308> NM_004994

<400> 165

```

agacacctgc gccctcacca tgagcctctg gcagccctg gtccctgggtgc tccctgggtgc 60
gggctgtctc ttgtgtccc ccagacagcg ccagtcacc cttgtgtctc tccctggaga 120
cctgagaacc aatctcacgg acaggcagct ggcagaggaa tacctgtacc gctatggtta 180
cactcggttg gcagagatgc gtggagagtc gaaatctctg gggcctgcgc tgctgcttct 240
ccagaagcaa ctgtccctgc ccgagaccgg tgagctggat agcgccacgc tgaaggccat 300
gcgaacccca cgggtcgggg tcccagacct gggcagattc caaacctttg agggcgacct 360
caagtggcac caccacaaca tcacctattg gatccaaaac tactcggaag acttgccgcg 420
ggcgggtgatt gacgacgcct ttgcccgcgc cttcgactg tggagcgcgg tgacgcgcgt 480
caccttcact cgcgtgtaca gccgggacgc agacatcgtc atccagtttg gtgtcgcgga 540
gcacggagac gggatatccct tcgacgggaa ggacgggctc ctggcacacg cctttcctcc 600
tggccccggc attcagggag acgcccattt cgacgatgac gagttgtggt cctggggcaa 660
gggcgtcgtg gttccaactc ggtttggaaa cgcagatggc gcggcctgcc acttcccctt 720
catcttogag ggccgtcctt actctgcctg caccaccgac ggtcgctccg acggttggcc 780
ctggtgcagt accacggcca actacgacac cgacgaccgg tttggcttct gccccagcga 840
gagactctac acccgggacg gcaatgctga tgggaaaccc tgccagtttc cattcatctt 900
ccaaggccaa tcctactccg cctgcaccac ggacggctgc tccgacggct accgctggtg 960
cgccaccacc gccaaactac accgggacaa gctcttcggc ttctgcccga cccgagctga 1020
ctcgacggtg atggggggca actcggcggg ggagctgtgc gtcttcccct tcaacttctt 1080
gggtaaggag tactcgacct gtaccagcga gggccgcgga gatgggccc tctggtgcgc 1140
taccacctcg aactttgaca gcgacaagaa gtggggcttc tgcccggacc aaggatacag 1200
tttgttcctc gtggcgcgcc atgagttcgg ccacgcgctg ggcttagatc attcctcagt 1260
gccggaggcg ctcatgtacc ctatgtaccg tctcactgag gggccccctt tgcataagga 1320
cgacgtgaat ggcatccggc acctctatgg tcttcgccct gaacctgagc cagggcctcc 1380
aaccaccacc acaccgcagc ccacggctcc cccgacggtc tgccccaccg gacccccacc 1440
tgtccacccc tcagagcgcc ccacagctgg cccacaggt cccccctcag ctggccccac 1500
aggtccccc actgctggcc cttctacggc cactactgtg cctttgagtc cgggtggacga 1560
tgctgcaaac gtgaacatct tcgacgccat cgcggagatt gggaaccagc tgtatttgtt 1620
caaggatggg aagtactggc gattctctga gggcaggggg agcggccgc agggcccctt 1680
ccttatcgcc gacaagtggc ccgcgctgcc ccgcaagctg gactcggctt ttgaggagcc 1740
gctctccaag aagcttttct tcttctctgg gcgccaggtg tgggtgtaca caggcgcgctc 1800
ggtgctgggc ccgagcgctc tggacaagct gggcctggga gccgacgtgg cccaggtgac 1860

```



```

cggggccctc cggagtgcca gggggaagat gctgctgttc agcggggcggc gcctctggag 1920
gttcgacgtg aaggcgagaga tgggtgatcc ccggagcgcc agcagagtggt accggatggt 1980
ccccggggtg cctttggaca cgcacgacgt cttccagtac cgagagaaaag cctattttctg 2040
ccaggaccgc ttctactggc gcgtgagttc ccggagtggg ttgaaccagg tggaccaagt 2100
gggctacgtg acctatgaca tcctgcagtg ccctgaggac tagggctccc gtctgtcttt 2160
gcagtgccat gtaaatcccc actgggacca accctgggga aggagccagt ttgccggata 2220
caaaactggtt ttctgttctg gaggaagggg aggagtggag gtgggctggg ccctctcttc 2280
tcacctttgt tttttgttgg agtgtttcta ataaacttgg attctctaac cttt 2334

```

```

<210> 166
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_004994

```

```

<400> 166
ggccctctct tctcaccttt gttttttgtt ggagtgtttc taataaactt ggattctcta 60

```

```

<210> 167
<211> 5329
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> Modified_base
<222> 1 ... 5329
<223> n = a,c,g, or t

```

```

<300>
<308> NM_005063

```

```

<400> 167
gtggtgtogg tgtcggcagc atccccggcg ccctgctgcg gtgcgcggag ccctcggcct 60
ctgttctcct cccctcccg cccttacctc cagcggggac cgcccgcgcc agtcaactcc 120
tcgcactttg cccctgcttg gcagcggata aaagggggct gaggaatac cggacacgtc 180
caccogttgc cagcttagc ctttaaatc ccggctcggg acctccacgc accgggctag 240
cgccgacaac cagctagcgt gcaaggcgcc gggctcagc gcgtaccggc gggcttcgaa 300
accgcagtc tccggcgacc ccgaactccg ctccggagcc tcagccccct ggaaagtgat 360
cccgcatcg gagagccaag atgccggccc acttgctgca ggacgatata tctagctcct 420
ataccaccac caccaccatt acagcgctc cctccagggt cctgcagaat ggaggagata 480
agttggagac gatgccccct tacttggaag acgacattcg ccctgatata aaagatgata 540
tatatgaccc cacctacaag gataaggaag gcccaagccc caaggttgaa tatgtctgga 600
gaaacatcat ccttatgtct ctgctacact tgggagccct gtatgggatc actttgatcc 660
ctacctgcaa gttctacacc tggctttggg ggtatttcta ctattttgtc agtgccctgg 720
gcataacagc aggagctcat cgtctgtgga gccaccgctc ttacaaagct cggctgcccc 780
tacggctctt tctgatcatt gccaacacaa tggcattcca gaatgatgtc tatgaatggg 840
ctcgtgacca ccgtgccac cacaagtttt cagaaacaca tgctgatcct cataattccc 900
gacgtggctt tttcttctct cagctgggtt ggctgcttgt gcgcaaacac ccagctgtca 960
aagagaaggg gactacgcta gacttgtctg acctagaagc tgagaaactg gtgatgttcc 1020
agaggaggtg ctacaaacct ggcttgctgc tgatgtgctt catcctgccc acgcttgctg 1080
cctgggtatt ctgggggtgaa acttttcaaa acagtgtgtt cgttgccact ttcttgcat 1140
atgctgtggt gcttaatgcc acctggctgg tgaacagtgc tgcccacctc ttcgatata 1200
gtccttatga caagaacatt agccccggg agaatatcct ggtttcactt ggagctgtgg 1260
gtgagggtt ccacaactac caccactcct ttccctatga ctactctgcc agtgagtacc 1320
gctggcacat caacttcacc acattcttca ttgattgcat ggccgcccctc ggtctggcct 1380
atgaccggaa gaaagtctcc aaggccgcca tcttgccag gattaaaaga accggagatg 1440
gaaactacaa gattggctga gtttggggtc cctcaggttc ctttttcaaa aaccagccag 1500
gcagagggtt taatgtctgt ttattaacta ctgaataatg ctaccaggat gctaaagatg 1560
atgatgttaa ccattccag tacagtattc ttttaaaatt caaaagtatt gaaagccaac 1620

```

aactctgcct	ttatgatgct	aagctgatat	tattttcttct	cttatccctct	ctctctttcta	1680
ggcccatgtg	cctccttttc	actttaatcg	ccctcctttc	ccttattgcc	tcccaggcaa	1740
gcagctggtc	agtctttgct	cagtgtccag	cttccaaagc	ctagacaacc	tttctgtagc	1800
ctaaaacgaa	tggcttttgc	tccagataac	tctctttcct	tgagctgttg	tgagctttga	1860
agtaggtggc	ttgagctaga	gataaaacag	aatcttcttg	gtagtccctc	gttgattatc	1920
ttcagcccag	gcttttgcta	gatggaatgg	aaaagcaact	tcatttgaca	caaagcttct	1980
aaagcnaggt	aaattgtcgg	gggagagagt	tagcatgtat	gaatgtaagg	atgagggaag	2040
cgaaggaacc	tctcgccatg	atcagacata	cagctgccta	cctaagtagg	acttcaagcc	2100
ccaccacata	gcatgcttcc	tttctctcct	ggctcggggg	aaaaagtggc	tgccgtgttt	2160
ggcaatgcta	attcaatgcc	gcaacatata	gttgaggccg	aggataaaga	aaagacattt	2220
taagtttgta	gtaaaagtgg	tctctgctgg	ggaagggttt	tcttttcttt	ttttctttta	2280
taacaaggag	atttcttagt	tcatatatca	agaagtcttg	aagttgggtg	tttccagaat	2340
tggtaaaaac	agcagctcat	agaattttga	gtattccatg	agctgctcat	tacagttctt	2400
tctcttttct	gctctgccat	cttcaggata	ttggttcttc	ccctcatagt	aataagatgg	2460
ctgtggcatt	tccaaacatc	caaaaaaagg	gaaggattta	aggaggtgaa	gtcgggtcaa	2520
aaataaaaata	tataacata	tatacattgc	ttagaacgtt	aaactattag	agtatttccc	2580
ttccaaagag	ggatgtttgg	aaaaaactct	gaaggagagg	aggaattagt	tgggatgcca	2640
atttctctct	cactgctgga	catgagatgg	agaggctgag	ggacaggatc	tataggcagc	2700
ttctaagagc	gaacttcaca	taggaaggga	tctgagaaca	cgttcagggg	ttgagaaggt	2760
tactgagtga	gttattggga	gtcttaataa	actagatatt	aggteccattc	attaattagt	2820
tcagttttct	ccttgaatag	agtaaaaaact	agaaggcttc	tctccacagt	gttgtgcccc	2880
ttcactcatt	tttttttgag	gagaaggggg	tctctgttaa	catctagcct	aaagtataca	2940
aactgcctgg	ggggcagggt	taggaatctc	ttcactaccc	tgattcttga	ttcctggctc	3000
tacctgtctt	gtcccttttc	tttgaccaga	ttttctctt	ccctgaacgt	tttcttcttt	3060
ccctggacag	gcagctcctt	ttgtgtgtat	tcagaggcag	tgatgacttg	ctgtccaggc	3120
agctccctcc	tgacacacaga	atgctcaggg	tcactgaacc	actgcttctc	ttttgaaagt	3180
agagctagct	gccactttca	cgtggcctcc	gcagtgtctc	cacctacacc	cctgtgctcc	3240
cctgccacac	tgatggctca	agacaaggct	ggcaaaccct	cccagaaaca	tctctggccc	3300
agaaagcctc	tctctccctc	cctctctcat	gagaagccaa	gcgctcatgt	tgagccagtg	3360
ggccagccac	agagcaaaaag	agggtttatt	ttcagtcccc	tctctctggg	tcagaaccag	3420
agggcatgct	gaatgcccc	tgcttacttg	gtgagggtgc	ccgcctgag	tcagtgtctt	3480
cagctggcag	tgcaatgctt	gtagaagtag	tgagaaacag	ttctcactgg	gaagaagcaa	3540
gggcaagaac	ccaagtgcct	cacctcgaaa	ggaggccctg	ttccctggag	tcagggtgaa	3600
ctgcaaagct	ttggctgaga	cctgggattt	gagataccac	aaaccctgct	gaacacagtg	3660
tctgttcagc	aaactaacca	gcattcccta	cagcctaggg	cagacaatag	tatagaagtc	3720
tggaaaaaaa	caaaaacaga	atttgagaac	cctggaccac	tctgtccct	gtagctcagt	3780
catcaaagca	gaagtctggc	tttgctctat	taagattgga	aatgtacact	accaaact	3840
cagtccactg	ttgagcccca	gtgctggaag	ggaggaaggc	ctttcttctg	tgtaatttgc	3900
gtagaggcta	cagggttag	cctggactaa	agggatcctt	gtctttgagc	tattcacctc	3960
agtagaaaag	gatctaaggg	aagatcactg	tagtttagtt	ctgttgacct	gtgcacctac	4020
cccttgaaa	tgtctgctgg	tatttctaat	tccacaggtc	atcagatgcc	tgcttgataa	4080
tataataaca	ataaaaacaa	ctttcacttc	ttcctattgt	aatcgtgtgc	catggatctg	4140
atctgtacca	tgacctaca	taaggctgga	tggcacctca	ggctgagggc	cccaatgtat	4200
gtgtggctgt	gggtgtgggt	gggagtgtgt	ctgctgagta	aggaacacga	ttttcaagat	4260
tctaaagctc	aattcaagtg	acacattaat	gataaaactca	gatctgatca	agagtcggga	4320
tttctaacag	tccttgcttt	gggggtgtgt	ctggcaactt	agctcagggtg	ccttacatct	4380
tttctaatac	cagtgttgca	tatgagcctg	ccctcactcc	ctctgcagaa	tccctttgca	4440
cctgagaccc	tactgaagtg	gctggtagaa	aaaggggcct	gagtggagga	ttatcagtat	4500
cacgatttgc	aggattccct	tctgggcttc	attctggaaa	cttttgtttag	ggctgctttt	4560
cttaagtgcc	cacatttgat	ggagggtgga	aataatttga	atgtatttga	tttataagtt	4620
tttttttttt	tttgggttaa	aagatggttg	tagcattttaa	aatggaaaat	tttctccttg	4680
gtttgctagt	atcttgggtg	tatttctctgt	aagtgtagct	caaatagggtc	atcatgaaag	4740
gttaaaaag	cgagggtggc	atgttatgtc	ggtggttgcc	agggcctcca	accactgtgc	4800
cactgacttg	ctgtgtgacc	ctgggcaagt	cacttaacta	taagggtgcct	cagttttcct	4860
tctgttaaaa	tggggataat	aatactgacc	tacctcaaag	ggcagttttg	aggcatgact	4920
aatgcttttt	agaaagcatt	ttgggatcct	tcagcacagg	aattctcaag	acctgagtat	4980
tttttataat	aggaatgtcc	accatgaact	tgatacgtcc	gtgtgtccca	gatgctgtca	5040
ttagtctata	tggttctcca	agaaactgaa	tgaatccatt	ggagaagcgg	tggataacta	5100
gccagacaaa	atttgagaat	acataaacia	cgcatttgcca	cggaaacata	cagaggatgc	5160
cttttctgtg	attgggtggg	attttttccc	tttttatgtg	ggatatagta	gttacttgtg	5220

acaagaataa ttttgaata atttctatta atatcaactc tgaagctaatt tgtactaatc 5280

tgagattgtg tttgttcata ataaaagtga agtgaatctg attgcaactg 5329

<210> 168

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005063

<400> 168

aataatgcta ccaggatgct aaagatgatg atgttaaccc attccagtac agtattcttt 60

<210> 169

<211> 634

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005101

<400> 169

cggtgagag	gcagcgaact	catctttgcc	agtacaggag	cttgtgccgt	ggcccacagc	60
ccacagccca	cagccatggg	ctgggacctg	acgggtgaaga	tgctggcggg	caacgaattc	120
cagggtgtccc	tgagcagctc	catgtcgggtg	tcagagctga	aggcgagat	caccagaag	180
attggcggtgc	acgccttcca	gcagcgtctg	gctgtccacc	cgagcgggtg	ggcgctgcag	240
gacaggggtcc	cccttgccag	ccagggcctg	ggccctggca	gcacggtcct	gctgggtggtg	300
gacaaatgcg	acgaacctct	gagcatcctg	gtgaggaata	acaagggccg	cagcagcacc	360
tacgaggtcc	ggctgacgca	gaccgtggcc	cacctgaagc	agcaagtga	cgggctggag	420
ggtgtgcagg	acgacctgtt	ctggctgacc	ttcgagggga	agcccctgga	ggaccagctc	480
ccgctggggg	agtacggcct	caagcccctg	agcaccgtgt	tcatgaatct	gcgcctgcgg	540
ggaggcggca	cagagcctgg	cgggcggagc	taagggcctc	caccagcatc	cgagcaggat	600
caagggcccg	aaataaaggc	tggtgtaaga	gaat			634

<210> 170

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005101

<400> 170

tggtggtgga caaatgcgac gaacctctga gcatcctggt gaggaataac aagggccgca 60

<210> 171

<211> 1339

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005139

<400> 171

gaattccgat	tagtgtgatc	tcagctcaag	gcaaagggtg	gatatcatgg	catctatctg	60
ggttgacac	cgaggaacag	taagagatta	tccagacttt	agcccatcag	tggatgctga	120
agctattcag	aaagcaatca	gaggaattgg	aactgatgag	aaaatgctca	tcagcattct	180
gactgagagg	tcaaatgcac	agcggcagct	gattgttaag	gaatatcaag	cagcatatgg	240
aaaggagctg	aaagatgact	tgaaggtga	tctctctggc	cactttgagc	atctcatggt	300

```

ggccctagtg actccaccag cagtctttga tgcaaagcag ctaaagaaat ccatgaaggg 360
cgcggaaca aacgaagatg ccttgattga aatcttaact accaggacaa gcaggcaaat 420
gaaggatata tctcaagcct attatacagt atacaagaag agtcttgag atgacattag 480
ttccgaaca tctggtgact tccggaagc tctgttgact ttggcagatg gcagaagaga 540
tgaaagtctg aaagtggatg agcatctggc caaacaagat gccagattc tctataaagc 600
tggtgagaac agatggggca cggatgaaga caaattcact gagatcctgt gtttaaggag 660
ctttcctcaa ttaaaactaa catttgatga atacagaaat atcagccaaa aggacattgt 720
ggacagcata aaaggagaat tatctgggca ttttgaagac ttactgttgg ccatagttaa 780
ttgtgtgagg aacacgccg cctttttagc cgaaagactg catcgagcct tgaagggtat 840
tggaactgat gaggttactc tgaaccgaat aatggtgtcc agatcagaaa ttgacctttt 900
ggacattcga acagagttca agaagcatta tggctattcc ctatattcag caattaaatc 960
ggatacttct ggagactatg aaatcacact cttaaaaatc tgtggtggag atgactgaac 1020
caagaagata atctccaaag gtccacgatg ggctttccca acagctccac cttacttctt 1080
ctcatactat ttaagagaac aagcaaatat aaacagcaac ttgtgttcct aacaggaatt 1140
ttcattgttc tataacaaca acaacaaaag cgattattat tttagagcat ctcatattata 1200
atgtagcagc tcataaatga aattgaaaat ggtattaaag atctgcaact actatccaac 1260
ttataattct gctttcaag ttaagaatct ttatagttct actccattaa atataaagca 1320
agataataaa acggaattc 1339

```

<210> 172

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005139

<400> 172

```

ttcagcaatt aaatcggata cttctggaga ctatgaaatc acactcttaa aaatctgtgg 60

```

<210> 173

<211> 1582

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005165

<400> 173

```

ccgagctgtg cttgtggctg cggctgctaa ctggctgcgc acaggagct gtcaccatgc 60
ctcactcgta cccagccctt tctgctgagc agaagaagga gttgtctgac attgccctgc 120
ggattgtagc cccgggcaaa ggcattcttg ctgcggatga gtctgtaggc agcatggcca 180
agcggctgag ccaaattggg gtggaaca caagagagaa ccgccggctg taccgccagg 240
tcctgttcag tgctgatgac cgtgtgaaaa agtgcattgg aggcgtcatt ttcttccatg 300
agaccctcta ccagaaagat gataatggtg ttcccttcgt ccgaaccatc caggataagg 360
gcatcgtcgt gggcatcaag gttgacaagg tctggtgcc tctagctggg actgatggag 420
aaaccaccac tcaagggtcg gatgggctct cagaacgctg tgccaatac aagaaggatg 480
gtgctgactt tgccaagtgg cgctgtgtgc tgaaaatcag tgagcgtaca ccctctgcac 540
ttgccattct ggagaacgcc aacgtgctgg cccgttatgc cagtatctgc cagcagaatg 600
gcattgtgcc tattgtggaa cctgaaatat tgccctgatg agaccacgac ctcaaactgt 660
gtcagtatgt tacagagaag gtcttggctg ctgtgtacaa ggccctgagt gaccatcatg 720
tatacctgga ggggaccctg ctcaagccca acatggtgac cccgggccat gcctgtccca 780
tcaagtatac cccagagag attgcatgg caactgtcac tgccctgcgt cgcactgtgc 840
cccagctgt cccaggagt accttctctg ctgggggtca gagcgaagaa gagcatcat 900
tcaacctcaa tgccatcaac cgctgcccc ttccccgacc ctgggcgctt accttctcct 960
atgggcgtgc cctgcaagcc tctgcaactc atgcctggcg agggcaacgg gacaatgctg 1020
gggctgccac tgaggagtgc atcaagcggg ctgaggtgaa tgggcttgca gccagggca 1080
agtatgaagg cagtggagaa gatggtggag cagcagcaca gtcactctac attgccaacc 1140
atgcctactg agtatccact ccataccaca gcccttgccc cagccatctg cacccccttt 1200
tgcttgtagt catggcagg gccaaatagc tatgcagagc agagatgcct tcacctggca 1260
ccaactgtgc ttctttctc tcttcccttc cctctctca ttgctgcacc tgggaccata 1320

```

```

ggatgggagg atagggagcc cctcatgact gagggcagaa gaaattgcta gaagtcagaa 1380
caggatggct gggctctccc ctacctcttc cagctccac aattttccca tgatgaggta 1440
gcttctccct gggctctcct tcttgctgc cctgtctcct gggatcagag ggtagtacag 1500
aagccctgac tcatgccttg agtacatacc atacagcaaa taaatggtag caaaacaaaa 1560
aaaaaaaaaa aaaaaaaaaa aa 1582

```

```

<210> 174
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005165

```

```

<400> 174
gagggtagta cagaagccct gactcatgcc ttgagtacat accatacagc aaataaatgg 60

```

```

<210> 175
<211> 451
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005213

```

```

<400> 175
acttccctgt tcactttggt tccagcatcc tgtccagcaa agaagcaatc agccaaaatg 60
atacctggag gcttatctga ggccaaaccc gccactccag aaatccagga gattgttgat 120
aagggttaaac cacagcttga agaaaaaaca aatgagactt atggaaaatt ggaagctgtg 180
cagtataaaa ctcaagttgt tgctggaaca aattactaca ttaaggtacg agcaggtgat 240
aataaatata tgcacttgaa agtattcaaa agtcttcccg gacaaaatga ggacttggtg 300
cttactggat accagggtga caaaaacaag gatgacgagc tgacgggctt ttagcagcat 360
gtaccocaaag tgttctgatt ccttcaactg gctactgagt catgatcctt gctgataaat 420
ataacatca ataaagaagc attcttttcc a 451

```

```

<210> 176
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005213

```

```

<400> 176
aactggctac tgagtcatga tccttgctga taaatataac catcaataaa gaagcattct 60

```

```

<210> 177
<211> 366
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005218

```

```

<400> 177
gtcagctcag cctccaaagg agccagcctc tcccagttc ctgaaatcct gagtgttgcc 60
tgccagtcgc catgagaact tcctaccttc tgctgtttac tctctgctta cttttgtctg 120
agatggcctc aggtggtaac tttctcacag gccttggcca cagatctgat cattacaatt 180
gcgtcagcag tggagggcaa tgtctctatt ctgcctgccc gatctttacc aaaattcaag 240
gcacctgtta cagaggggaag gccaaagtgc gcaagtgagc tgggagtgac cagaagaaat 300
gacgcagaag tgaaatgaac tttttataag cattctttta ataaaggaaa attgcttttg 360

```

aagtat 366

<210> 178

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005218

<400> 178

gggagtgacc agaagaaatg acgcagaagt gaaatgaact ttttataagc attcttttaa 60

<210> 179

<211> 1519

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005326

<400> 179

```

ctgcctcggg acgctgtccc ccgcagcgac ggcccgttcc acctcgcgat ctgccgggta 60
cccgggcggg gtggcgctcg gcctccaggg atccactgtg cggtgccaaa aaagaggcgg 120
aggctcggcg cacagctctc ccggcgacgc tctcgggccc ccgccgccgc tcccaggccc 180
gtctcccgcc ccgtggcagt cggggctcgc ggacaaaaca agttgagcgc gagcgcggtg 240
attggttggc ggacggtgcg aggtggacgc tgattggctg agggcagcgc gaggcgggcg 300
ctgattggct gcgacgcgcc gacgccggtg ttttgcaatc ctgggcagct cggcagtgca 360
gcccggcccc ggtcatggtg gtgggcccag ggctgctcgg ccgccgcagc ctgcgccgcg 420
tgggagccgc ctgcgccgcg cgaggcctcg gtccagccct gctgggagtt ttctgccaca 480
cagatttgcg gaagaacctg accgtggacg agggcaccat gaaggtagag gtgctgcctg 540
ccctgaccga caactacatg tacctggtca ttgatgatga gaccaaggag gctgccattg 600
tggatccggt gcagccccag aaggtcgtgg acgcggcgag aaagcacggg gtgaaactga 660
ccacagtgtc caccaccac caccactggg accatgctgg cgggaatgag aaactggtca 720
agctggagtc gggactgaag gtgtacggg gtgacgaccg tatcggggcc ctgactcaca 780
agatcactca cctgtccaca ctgcaggtgg ggtctctgaa cgtaagtgct ctggcgaccc 840
cgtgccacac ttcaggacac atttgttact tcgtgagcaa gcccgaggc tcggagcccc 900
ctgccgtggt cacaggtgac acctgtttg tggctggctg cgggaagtgc tatgaaggga 960
ctgcggatga gatgtgtaaa gctctgctgg aggtcttggg ccgggtcccc ccggacacaa 1020
gagtctactg tggccacgag tacaccatca acaacctcaa gtttgacgc cacgtggagc 1080
ccggcaatgc cgccatccgg gagaagctgg cctgggcca ggagaagtac agcatcgggg 1140
agccacacgt gccatccacc ctggcagagg agtttaccta caacccttc atgagagtga 1200
gggagaagac ggtgcagcag cacgcaggtg agacggaccc ggtgaccacc atgcggggccg 1260
tgcgaggagg gaaggaccag ttcaagatgc cccgggactg aggcgcgccct gcaccttcag 1320
cggatttggg gattaggctc ttttaggtaa ctggctttcc tgctggtccg tgcgggaaat 1380
tcagtcttga tttaacctta attttacagc ccttggtctg tggtatcgga cattctaag 1440
cataattata agagaagttt aacaagtatt tattcccata aaaaaaaaaa aaaaaaaaaa 1500
aaaaaaaaaa aaaaaaaaaa 1519

```

<210> 180

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005326

<400> 180

cttgtgttat cggacattct aatgcatatt tataagagaa gtttaacaag tattttattcc 60

<210> 181

<211> 3378

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005461

<400> 181

acagctgcac	cgccgagctg	cgagcggctg	cgagcgagag	agcgtaagag	caagagagct	60
agagagcgag	caacgggcac	tcgccccacg	cctccccca	gccccaccgc	gcgctccgct	120
tgctctcca	ccccgcccga	ctctaccggg	cccggtccct	gcgcgggcac	agcccagagc	180
tctggggcgg	tgacggcagc	ctcgggactc	tccggcgcg	cgccgcgtcc	ccagacaaag	240
gcttggccgg	cgcccccggc	ccgctgcgcc	ctcgctcccc	gcctccccag	ctcttctccg	300
ctcttcccc	ccgcgcttgg	ctcggcgcg	tccggccggc	cgcaaagttt	cccgggcggc	360
agcgcgggct	gcgcctcgct	tcagcgatgg	ccgcgagct	gagcatgggg	ccagagctgc	420
ccaccagccc	gctggccatg	gagtatgtca	acgacttcga	cctgctcaag	ttcgacgtga	480
agaaggagcc	actggggcgc	gcggagcgtc	cgggcaggcc	ctgcacacgc	ctgcagccag	540
ccgctcgggt	gtcctccaca	ccgctcagca	ctccgtgtag	ctccgtgcc	tcgtcgccca	600
gcttcagccc	gaccgaacag	aagacacacc	tcgaggatct	gtactggatg	gcgagcaact	660
accagcagat	gaaccccag	gcgctcaacc	tgacgcccga	ggacgcgggtg	gaagcgctca	720
tcggtctgca	cccagtgcc	cagccgctgc	aaagcttcga	cagctttcgc	ggcgctcacc	780
accaccacca	tcaccaccac	cctcaccgc	accacgcgt	cccggggccc	ggcggtggcc	840
acgacgagct	gggcccgcac	gctcaccgc	accatcacca	tcatacacia	gcgtcgccgc	900
cgccgtccag	cgccgctagc	ccggcgcaac	agctgcccac	tagccacccc	gggcccgggc	960
cgcacgcgac	ggcctcggcg	acggcgcgcg	gcggcaacgg	cagcggtggag	gaccgcttct	1020
ccgacgacca	gctcgtgtcc	atgtccgtgc	gcgagctgaa	ccgccacctg	cggggcttca	1080
ccaaggacga	ggtgatccgc	ctgaagcaga	agcgggcgac	cctgaagaac	cggggctacg	1140
cccagtcttg	caggtataaa	cgcgctccagc	agaagcacca	cctggagaat	gagaagacgc	1200
agctcattca	gcaggtggag	cagcttaagc	aggaggtgtc	ccggctggcc	cgcgagagag	1260
acgcctacaa	ggtcaagtgc	gagaaactcg	ccaactccgg	cttcagggag	gcgggctcca	1320
ccagcgacag	cccctcctct	cccagttct	ttctgtgagt	cgtggccgggt	cctggccccc	1380
gcccttgccc	cgccccggac	tccctgtccc	acgtccctag	tcccagacta	ccccggaccc	1440
tgtccctgcc	gcggccccag	ccttgacctg	tttgacttga	gcgagaggga	ggaaggcgcc	1500
gcgggcccgc	ggcgacgggc	gggtgcgagg	gcgggcaggg	gaccttggt	aaggcgagag	1560
tagcgcaogc	cagcgccggc	tcctagactc	gagcagagcc	ggagagagag	acgagagggt	1620
gggaggtooc	ggagtaactt	ctctccaggc	tgaaggcgcg	cgaggcatag	tcccagagaag	1680
tcaccaaggc	catctggaga	ctcctggctt	tctgaacttt	gcgcgttaag	ccgggacagc	1740
tgctttgctg	cccggagagt	agtcgcgcgc	aggaagagag	caacgaggaa	aggagaggga	1800
ctctggcgtc	ccggcaggcg	agaggcgagg	ctgagcgaaa	gaagggaagg	cagacggacc	1860
tgtctgtcag	agttcggaga	acactggctc	tcagccctga	gacacaggcc	tcagttagga	1920
cgctcggcgc	ccaaatctca	tcagttttat	tgctgtctcg	attatataga	aaaatacaaa	1980
aaatctgcat	taaaaaatatt	aatcctgcat	gctggacatg	tatggtaata	atttctattt	2040
tgtaccattt	tcttgtttaa	cttttagcatg	ttgttgatca	tggatcatac	tcccttgttt	2100
tctttgggtg	agaagggatc	gcagtttgga	aaactccggc	gctgcgtgcg	gggtttcagt	2160
cccagctgta	ggcttgtaaa	taccgcccc	gccaaccgc	atagagaacg	tggcagcaag	2220
ctgagggctc	ttgtttgggt	ttattattac	ggatattttg	tttgtaagtt	aaaaagaaaa	2280
aaaaaaagaa	aaagttccgg	gcattttgca	tcagaaaaca	actttgtctt	ggggcacact	2340
tggaagttgc	atgttttctt	tccttccctt	atccccattc	ggctcctctt	ttcctctctc	2400
gctttagttt	tcaaccttgt	tggtgctgag	agagagaacc	gagaggctcc	agtacaaggg	2460
cagggcaggg	cagggaagct	gccaaagctc	gcaccccaga	ggagtgttct	ggactacagc	2520
cttgccttat	ggtcaaattg	atacccttaa	taagaaagga	aaggaaagga	aaacagatcc	2580
tcccctctga	tttttattgt	aaccagaatc	accctgaggt	cccttctgaa	ccctctgggc	2640
ctgcgctaag	tgtaggagcc	acagcgctcc	taggttgaga	ggcttagcca	tcctgacccc	2700
tggcagtga	ctggtaagca	gacactgcac	tgaaccaact	gctatgctca	gaatgtacca	2760
gaaacccaaa	cattggcaag	taattttgca	actttcaagt	gcgttcttta	gaccaatgca	2820
ttgcgtttct	ttccctgctt	ttgagatagt	aggaagagtt	cttgggtggtg	tccccccct	2880
tcaattcttc	agttgtatag	tagttatagg	gaagatatgg	gtgtttttct	ttattattac	2940
tttttttttt	ctgcaggtca	gtaaaaggat	ttaagttgca	ctgacaaaaa	tacaaaaata	3000
aaagtgtatt	tttaagttcc	cattttgaaat	tgctggcgct	gctggccgga	tgcatttttg	3060
agtttgtatt	agttgtataa	ttaacagtaa	taacaagatt	gtatgaaccg	catggtgctt	3120
gcagttttta	atattgtgga	tatttgcctt	gcatacagaaa	cgagcttttg	tttttacaga	3180

```

ttcaactgtg ttgaaatcaa acctgccgca acagaaattg tttttatttc atgtaaaata 3240
agggatcaat ttcaaaccct gcttatgata tgaaaatatt aaaacctagt ctattgtagt 3300
tttattcaga ctggtttctg ttttttggtt attaaaatgg tttcctattt tgcttattaa 3360
aaaaaaaaa aaaaaaaa 3378

```

```

<210> 182
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005461

```

```

<400> 182
atttgctctg catcagaaac gagctttggt ttttacagat tcaactgtgt tgaaatcaaa 60

```

```

<210> 183
<211> 597
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005532

```

```

<400> 183
agctgaagtt gaggatctct tactctctaa gccacggaat taacccgagc aggcattggag 60
gcctctgctc tcacctcatc agcagtgacc agtgtggcca aagtggtcag ggtggcctct 120
ggctctgccg tagttttgcc cctggccagg attgctacag ttgtgattgg aggagttgtg 180
gccatggcgg ctgtgcccat ggtgctcagt gccatgggct tcaactgggc gggaatcgcc 240
tcgtcctcca tagcagccaa gatgatgtcc gcggcggcca ttgccaatgg ggggtggagtt 300
gcctcgggca gccttggtgg tactctgcag tcaactggag caactggact ctccggattg 360
accaagttca tctctgggctc cattgggtct gccattgcgg ctgtcattgc gaggttctac 420
tagtcctctg cccctcgccc tgcagagaag agaaccatgc caggggagaa ggcaaccagc 480
catcctgacc cagcgaggag ccaactatcc caaatatacc tgggtgaaat ataccaaatt 540
ctgcatctcc agaggaaaat aagaaataaa gatgaattgt tgcaactctt aaaaaaa 597

```

```

<210> 184
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005532

```

```

<400> 184
agccaactat cccaaatata cctgggtgaa atataccaaa ttctgcatct ccagaggaaa 60

```

```

<210> 185
<211> 1661
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_005566

```

```

<400> 185
tgctgcagcc gctgccgccg attccggatc tcattgccac gcgccccga cgaccgcccg 60
acgtgcattc ccgattcctt ttggttccaa gtccaatatg gcaactctaa aggatcagct 120
gatttataat cttctaaagg aagaacagac cccccagaat aagattacag ttgttggggt 180
tgggtgctgt ggcatggcct gtgccatcag tatcttaatg aaggacttgg cagatgaact 240
tgctcttggt gatgtcatcg aagacaaatt gaaggagag atgatggatc tccaacatgg 300

```



```

cagccttttc cttagaacac caaagattgt ctctggcaaa gactataatg taactgcaaa 360
ctccaagctg gt cattatca cggctggggc acgtcagcaa gagggagaaa gccgtcttaa 420
tttgggtccag cgtaacgtga acatatttaa attcatcatt cctaattgtg taaaatacag 480
cccgaactgc aagttgctta ttgtttcaaa tccagtggat atcttgacct acgtggcttg 540
gaagataagt ggttttccca aaaaccgtgt tattggaagt ggttgcaatc tggattcagc 600
ccgattccgt taacctgatgg gggaaaggct gggagttcac ccattaagct gtcattgggtg 660
ggtccttggg gaacatggag attccagtgt gcctgtatgg agtggaatga atgttgctgg 720
tgtctctctg aagactctgc acccagattt agggactgat aaagataagg aacagtggaa 780
agaggttcac aagcaggtgg ttgagagtgc ttatgaggtg atcaaaactca aaggctacac 840
atcctgggct attggactct ctgtagcaga tttggcagag agtataatga agaactcttag 900
gcggtgacac ccagtttcca ccatgattaa gggctcttac ggaataaagg atgatgtctt 960
ccttagtggt ccttgcatct tgggacagaa tggaatctca gaccttgatg aggtgactct 1020
gacttctgag gaagaggccc gtttgaagaa gagtgcagat acactttggg ggatccaaaa 1080
ggagctgcaa ttttaaagtc ttctgatgtc atatcatttc actgtctagg ctacaacagg 1140
attctaggtg gaggttggtg atgttgctct ttttatctga tctgtgatta aagcagtaat 1200
attttaagat ggactgggaa aaacatcaac tctgaagtt agaaataaga atgggttgta 1260
aatccacag ctatatcctg atgctggatg gtattaatct tgtgtagtct tcaactgggt 1320
agtgtgaaat agttctgcca cctctgacgc accactgcca atgctgtacg tactgcattt 1380
gccccttgag ccaggtggat gtttaccgtg tggtatataa cttcctggct ccttactga 1440
acatgcctag tccaacattt tttccagtg agtcacatcc tgggatccag tgtataaate 1500
caatatcatg tcttgtgcat aattcttcca aaggatctta ttttgtgaac tatatcagta 1560
gtgtacatta ccatataatg taaaaagatc tacatacaaa caatgcaacc aactatccaa 1620
gtgttatacc aactaaaacc ccaataaac cttgaacagt g 1661

```

<210> 186

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005566

<400> 186

```

catcaactcc tgaagttaga aataagaatg gtttgtaaaa tccacagcta tatcctgatg 60

```

<210> 187

<211> 2993

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005689

<400> 187

```

gggcctgcag ttggcagaag ggtcccgggc ccagagccag cggggccgtg ctgagacggc 60
gtacgtgccc tgctgagtg cgtggcggcg gcgcgtgcgc taggggagtg ggcggtgagg 120
cctgggtccac gtgcgtccct tcccgggacc ccgcagctt ggcgcccagc ggctacgtga 180
gccaaggcac ccgatgtcc gcgcccctct ccgagtgaac agtcccggcc tccgggtccc 240
cagtgcgccg agcctcgggc ggcgtccacg cattgccatg gtgactgtgg gcaactactg 300
cgaggccgaa gggcccgtgg gtccggcctg gatgcaggat ggccctgagtc cctgcttctt 360
cttcacgctc gtgcccctga cgcgatggc tctagggact ctggccttgg tgctggctct 420
tccctgcaga cgcggggagc ggcccgtgg tgctgattcg ctgtcttggg gggcggggcc 480
tcgcatctct ccctacgtgc tgcagctgct tctggccaca cttcaggcgg cgtgccccct 540
ggcgggcttg gctggccggg tgggactgc ccggggggcc ccactgcca gctatctact 600
tctggcctcc gtgtgggaga gtctggccgg cgcctgtggc ctgtggctgc ttgtcgtgga 660
gcggagccag gcacggcagc gtctggcaat gggcatctgg atcaagttca ggacacagccc 720
tggtctcctg ctctctgga ctgtggcgtt tgcagctgag aacttggccc tgggtgtctt 780
gaacagccca cagtgggtgg gggcaagggc agacttgggc caacaggttc agtttagcct 840
gtgggtgctg cggtatgtgg tctctggagg gctgttctct ctgggtctct ggccccttg 900
acttcgtccc cagtctata cattgcaggt tcatgaagag gaccaagatg tggaaaggag 960
ccaggttcgg tcagcagccc aacagtctac ctggcgagat tttggcagga agctccgcct 1020

```

```

cctgagtggc tacctgtggc ctogagggag tccagctctg cagctggtgg tgctcatctg 1080
cctggggctc atgggttttg aacgggcact caatgtgttg gtgcctatat tctataggaa 1140
cattgtgaac ttgctgactg agaaggcacc ttggaactct ctggcctgga ctgttaccag 1200
ttacgtcttc ctcaagttcc tccagggggg tggcactggc agtacaggct tcgtgagcaa 1260
cctgcgcacc ttctgttgga tccgggtgca gcagttcacg tctcggcggg tggagctgct 1320
catcttctcc cacctgcacg agctctcact gcgctggcac ctggggcgcc gcacagggga 1380
gggtgctgcg atcgcgatc ggggcacatc cagtgtcaca gggctgctca gctacctggt 1440
gttcaatgtc atccccacgc tggccgacat catcattggc atcatctact tcagcatggt 1500
cttcaacgcc tgggtttggc tcattgtgtt cctgtgcacg agtctttacc tcacctgac 1560
cattgtggtc actgagtga gaaccaagtt tcgtcgtgct atgaacacac aggagaacgc 1620
taccggggca cgagcagtgg actctctgct aaacttcgag acggtgaagt attacaacgc 1680
cgagagttaac gaagtggaaac gctatcgaga ggccatcatc aaatatcagg gtttgagtg 1740
gaagtcgagc gcttactagg ttttactaaa tcagaccag aacctggtga ttgggctcgg 1800
gctcctcgcc ggctccctgc tttgcgcata ctttgtcact gagcagaagc tacaggttgg 1860
ggactatgtg ctctttggca cctacattat ccagctgtac atgcccctca attggtttgg 1920
cacctactac aggatgatcc agaccaactt cattgacatg gagaacatgt ttgacttgct 1980
gaaagaggag acagaagtga aggaccttcc tggagcaggg ccccttcgct ttcagaaggg 2040
ccgtattgag tttgagaacg tgcaacttcag ctatgccgat gggcgggaga ctctgcagga 2100
cgtgtctttc actgtgatgc ctggacagac acttgccctg gtgggcccac ctggggcagg 2160
gaagagcaca attttgcgcc tgctgtttcg cttctacgac atcagctctg gctgcatccg 2220
aatagatggg caggacattt cacaggtgac ccaggcctct ctccggtctc acattggagt 2280
tgtgccccaa gacactgtcc tctttaatga caccatcgcc gacaatatcc gttacggccg 2340
tgtcacagct gggaatgatg aggtggaggc tgctgctcag gctgcaggca tccatgatgc 2400
cattatggct ttccctgaag ggtacaggac acagggtggc gagcggggac tgaagctgag 2460
cggcggggag aagcagcgcg tcgccattgc ccgcaccatc ctcaaggctc cgggcatcat 2520
tctgtctgat gaggcaacgt cagcgtgga tacatctaag gagagggcca tccaggcttc 2580
tctggccaaa gtctgtgcca accgcaccac catcgtagtg gcacacaggc tctcaactgt 2640
ggatcaatgct gaccagatcc tcgtcatcaa ggatggctgc atcgtggaga ggggacgaca 2700
cgaggctctg ttgtcccgag gtgggtgtga tgctgacatg tggcagctgc agcagggaca 2760
ggaagaaacc tctgaagaca ctaagcctca gaccatggaa cgggtgacaaa agtttggcca 2820
cttccctctc aaagactaac ccagaaggga ataagatgtg tctcctttcc ctggcttatt 2880
tcatcctggt cttggggtat ggtgctagct atggtaaggg aaagggacct ttccgaaaaa 2940
catcttttgg ggaaataaaa atgtggactg tgaaaaaaa aaaaaaaa aaa 2993

```

<210> 188
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005689

<400> 188
 ggaaaggagc ctttccgaaa aacatctttt ggggaaataa aaatgtggac tgtgaaaaaa 60

<210> 189
 <211> 1830
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005749

<400> 189
 ggggagttga aacctaattt tgtggcgtag cagctatgca gcttgaaatc caagtagcac 60
 taaattttat tatttcgtat ttgtacaata agcttcccag gagacgtgtc aacatttttg 120
 gtgaagaact tgaagactt cttaagaaga aatatgaagg gcactgggtat cctgaaaagc 180
 catacaaagg atcggggttt agatgtatac acatagggga gaaagtggac ccagtgtatt 240
 aacaagcatc caaagagagt ggttttgaca ttgatgatgt tcgtggcaat ctgccacagg 300
 atcttagtgt ttggatcgac ccatttgagg tttcttacca aattggtgaa aagggaccag 360
 tgaagggtgct ttacgtggat gataataatg aaaatggatg tgagttggat aaggagatca 420

```

aaaacagctt taaccagag gccaggtt ttatgccat aagtgacca gcctcatcag 480
tgtccagctc tccatcgct ccttttggtc actctgctgc tgtaagccct accttcacgc 540
cccgggtccac tcagccttta acctttacca ctgccacttt tgctgccacc aagttcggt 600
ctaccaaagt gaagaatagt ggccgtagca acaaggttgc acgtacttct cccatcaacc 660
tcggccttgaa tgtgaatgac ctcttgaagc agaaagccat ctcttcctca atgcactctc 720
tgtatgggct tggcttgggt agccagcagc agccacagca acagcagcag ccagcccagc 780
cgccaccgcc accaccacca ccacagcagc aacaacagca gaaaacctct gctctttctc 840
ctaattgcaa ggaatttatt tttcctaata tgcagggtca aggtagtagt accaatggaa 900
tgttccagg tgacagcccc cttaacctca gtcctctcca gtacagtaat gcctttgatg 960
tgtttgcagc ttaattgggc ctcaatgaga agtcttttgt agatggcttg aatttttagt 1020
taaataacat gcagatttct aaccagcaat tccagcctgt tatggctaac taaaaaaaaa 1080
aaaatgtatc gtacaagtta aaatgcacgg gcccaagggg gatTTTTTTT ttcacctctc 1140
tgagaatttt tttttttaag cttatagtaa ggatacatc aagcttggtt aaaaaataa 1200
taataaaaca tgcattttt ttcatTTTg ccacaaagc aaagttattt tatactgact 1260
gtatatTTta aagtatactc tcagatatgg cctcttacag tatttaagat atagcaagga 1320
catggctgat ttttttttat aaaaattggc actaataagt gggtttattg gtcttttcta 1380
attgtataat ttaatttagt acaaagttg taaaatatca gaggatata atatatTgtt 1440
tctacgacat ggtattgcac ttatatcttt ttactacag gatctgtgac agcagcagct 1500
tcatgttgta ttttttttac tgaaattgta aaatatccat cttaaagaca tcaactattc 1560
taaaaattgt gtacaggata ttcttttagt ggtggaatta aaatgtacga atacttgctt 1620
tttcaaaaaa atgtattttc tgttaaaagt ttaaagattt ttgctatata ttatggaaga 1680
aaaatgtaat cgtaaatatt aattttgtac ctatattgtg caatacttga aaaaaacggt 1740
ataaaagtat tttgagtcag tgtcttacat gttaagaggg actgaaatag tttatatata 1800
gtttgtatta aaattcttta aaattaaaaa 1830

```

<210> 190

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005749

<400> 190

```

aaacctctgc tctttctcct aatgccaaagg aatttatttt tcctaatatg cagggtcaag 60

```

<210> 191

<211> 1534

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005804

<400> 191

```

ggaagcgcag caactcgtgt ctgagcgccc ggcggaaaac cgaagttgga agtgtctctt 60
agcagcgcgc ggagaagaac ggggagccag catcatggca gaacaggatg tggaaaacga 120
tcttttgat tacgatgaag aggaagagcc ccaggctcct caagagagca caccagctcc 180
ccctaagaaa gacatcaagg gatcctacgt ttccatccac agctctgggt tccgggactt 240
tctgctgaag ccggagctcc tgcgggccat cgtggactgt ggctttgagc atccttctga 300
ggtccagcat gagtgcattc ccaggccat cctgggcatg gacgtcctgt gccaggccaa 360
gtccgggatg ggcaagacag cggctcttct gtctggccacc ctacagcaga ttgagcctgt 420
caacggacag gtgacggtcc tggtcattgt ccacacgagg gagctggcct tccagatcag 480
caaggaatat gagcgctttt ccaagtacat gccacgcgtc aaggtgtctg tgttcttcgg 540
tggctctctc atcaagaagg atgaagaagt gttgaagaag aactgtcccc atgtcgtggt 600
ggggaccccg ggccgcatcc tggcgctcgt gcggaatagg agcttcagcc taaagaatgt 660
gaagcacttt gtgctggacg agtgtgacaa gatgctggag cagctggaca tgcggcggga 720
tgtgcaggag atcttcggcc tgacaccaca cgagaagcag tgcattgatg tcagcgccac 780
cctgagcaag gacatccggc ctgtgtgcag gaagttcatg caggatccca tggagggtgt 840
tgtggacgac gagaccaagg tcacgctgca cggcctgcag cagtactacg tcaactcaa 900

```

```

agacagtgag aagaaccgca agctctttga tctcttggat gtgctggagt ttaaccaggt 960
gataatcttc gtcaagtcag tgcagcgtg catggccctg gccagctcc tegtggagca 1020
gaacttcccg gccatcgcca tccaccggg catggcccag gaggagcgcc tgtcacgcta 1080
tcagcagttc aaggatttcc agcggcggat cctggtggcc accaatctgt ttggccgggg 1140
gatggacatc gagcgagtca acatcgtctt taactacgac atgcctgagg actcggacac 1200
ctacctgcac cgggtggccc gggcgggtog ctttggcacc aaaggcctag ccatcacttt 1260
tgtgtctgac gagaatgatg ccaaaatcct caatgacgtc caggaccggt ttgaagttaa 1320
tgtggcagaa ctccagagg aaatcgacat ctccacatac atcgagcaga gccggtaacc 1380
accacgtgcc agagccgccc acccggagcc gcccgcatgc agcttcacct cccctttcca 1440
ggcgccactg ttgagaagct agagattgta tgagaataaa cttgttatta tggaaaaaaa 1500
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 1534

```

<210> 192
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005804

<400> 192
 gttgagaagc tagagattgt atgagaataa agtgttatta tgaaatgaag aagcctcacc 60

<210> 193
 <211> 1416
 <212> DNA
 <213> Homo sapiens

<220>
 <221> Modified_base
 <222> 1 ... 1416
 <223> n = a,c,g, or t

<300>
 <308> NM_005945

<400> 193
 aggaattccg gaattccgga attccgatgg atggaacaga aaataaatct aagtttggtg 60
 cgaacgccat tctgggggtg tcccttgccg tctgcaaagc tgggtgccgtt gagaaggggg 120
 tcccctgtac cgccacatcg cgtacttggc tggcaacttc gaagtcaccc tgccagtcac 180
 ggcgttcaag tgtcatcatc aatggcgggt ctcatgtctg caacaagctg gccatgcaga 240
 gtctgtcttc ccagtcggtg cagcaaacctc aggggaagcca tgccgcattg gaggagaggt 300
 ttaccacaac ctgaagaatg tcatcaagga gaaatatggg aaagatgccca ccaatgtggg 360
 gatttgccgc ggtttgctcc caacatcctg gagaataaag aaggcctgga gctgctgaag 420
 actgctattg gaaagcctgg cctacactgt aaaggtggtc atggcatgga cgtagcggcc 480
 tccgagttct tcaggtcagg gaactatgac ctggacttca agtctccga tgacccagc 540
 aggtacatct cgcccgacca gctggctgac ctgtacaagt ccttcatcaa ggactaccca 600
 gtggtgtcta tcgaagatcc ctttgaccag gatgactggg gagcttcaga agttcacagc 660
 cagtgcagga atccaggtag tgggggggatg actcacagt accaaccaca agaggatcgc 720
 caagggcgtga acgagaagtc ctgcaactgc ctccctgctca aagtcaacca gattggctcc 780
 gtgaccgagt ctcttcaggc gtgcaagctg gccaggcca atgggtgggg cgctcatggtg 840
 tctcatcggt cggggggagac tgaagatacc ttcacgcgtg acctgggtgt ggggctgtgc 900
 actggggcag atcaagactg gtgcccttg ccatcacgc gcttggccaa gtacaaccag 960
 ctcttcagaa ttgaagagga gctgggcagg aaggctaagt ttgccggcag gaacttcaga 1020
 aacccttggt ccaagtaagc tgtgggcagg caagccttcg gtacactgtt ggctacagac 1080
 ccctcccttg gtgtcagctc aggcagctcg agggcccgga ccaacacttg caggggtccc 1140
 tgctagttag cggccaccgc cgtggagtgc gtaccgcttc cttagaactc tacagaagcc 1200
 aagctccctg gaagccctgt tggcagctct agctttgcag ttgtgtaatt ggcccaagtc 1260
 attgtttttc tcgccttact ttccaccaag tgtctagagt catgtgagcc tngtgtcate 1320
 tccgggggtg ccacaggcta gatcccggt ggttttgcgc tcaaaataaa aagcctcagt 1380
 gaccatgaa aaaaaaaaaa gaattccgga attccg 1416

<210> 194
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_005945

<400> 194
 ttgtgtaatt ggcccaagtc attgtttttc tcgccttact ttccaccaag tgtotagagt 60

<210> 195
 <211> 961
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006014

<400> 195
 ggcgaccacg gtgtcttcaa aagccccgtc agggttggct tcctggggcc ggaccgactg 60
 tgggtcagtt tgcaccagcg ctctggaatc gagttacgcg cgaaagggca gagtttctgg 120
 aggaaaccgc agcctctcaa ccgctgaccg ggtctcagaa ggcccccgcc agggccgctt 180
 ggcgggaact gaccacgcgc cagtcaggct ctccaggac ctgcgaggc gcgtgtgggc 240
 ggagtcgtgc gcagggggcg gggcttcggg aaggagccac agagagggcg gggcgtagga 300
 cctgcgcttc ggggtggag tcggagcggc gcggcgggcg tcatgcggga cgcggatgca 360
 gacgcaggcg gaggcgctga cggcggggat ggccgggggtg gccacagctg ccgcgggggc 420
 gtggacacag ccgcagctcc ggccgggtgga gtcccccag cgcacgcgcc aggtccgggc 480
 agagacgccg cgtctgcggc caggggggtca cgaatgcggc cgcacatatt caccctcagc 540
 gtgcctttcc cgacccctt ggaggcggaa atcgcccatg ggtccctggc accagatgcc 600
 gagccccacc aaagggtggt tgggaaggat ctacacagtga gtggcaggat cctggtcgtc 660
 cgctggaag ctgaagactg tcgcctgctc cgaatttccg tcatcaactt tcttgaccag 720
 cttccctgg tggcgcgac catgcagcgc tttggggccc ccgtttccc ctaagcctgg 780
 cctgggcaaa tggagcgagg tcccactttg cgtctccttg taggcagtgc gtccatcctt 840
 ccctagggca ggaattccca cagttgctac tttcctggga gggcctcatg ttttatctgg 900
 ttcttaaatg tttgttacta cagaaaataa aactgcgcta ctaaaaaaaa aaaaaaaaaa 960
 a 961

<210> 196
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006014

<400> 196
 ggcctcatgt tttatctggt tottaaatgt ttgttactac agaaaataaa actgaggtat 60

<210> 197
 <211> 1648
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006086

<400> 197
 atgcgggaga tcgtgcacat ccaggccggc cagtgcggca accagatcgg ggccaagttc 60
 tgggaagtca tcagtgatga gcatggcatc gacccacgag gcaactacgt gggcgactcg 120

```

gacttgcagc tggagcggat cagcgtctac tacaacgagg cctcttctca caagtacgtg 180
cctcgagcca ttctggtgga cctggaaccc ggaaccatgg acagtgtccg ctgaggggcc 240
tttggacatc tcttcaggcc tgacaatttc atctttggtc agagtggggc cggcaacaac 300
tgggccaagg gtactacac ggagggggcg gagctggtgg attcggtcct ggatgtggtg 360
cgggaaggagt gtgaaaactg cgactgcctg cagggcttcc agctgaccca ctcgctgggg 420
ggggggacgg gctccggcat gggcacgttg ctcatcagca aggtgcgtga ggagtatccc 480
gaccgcatca tgaacacctt cagcgtcgtg ccctcaccca aggtgtcaga cagggtggtg 540
gaacctaca acgccacgct gtccatccac cagctggtgg aaaacacgga tgaacctac 600
tgcacgcaca acgagggcgt ctacgacatc tgcttccgca ccctcaagct ggccacgccc 660
acctacgggg acctcaacca cctggtatcg gccaccatga gcggagtcac cactccttg 720
cgcttccggg gccagctcaa cgctgacctg cgcaagctgg ccgtcaacat ggtgcccttc 780
ccgcgcctgc acttcttcat gcccgcttc gccccctca ccaggcgggg cagccagcag 840
taccggggcc tgaccgtgcc cgagctcacc cagcagatgt tcgatgccaa gaacatgatg 900
gccgcctgcg accgcgcga cgccgcctac ctgacggtgg ccaccgtgt ccggggccgc 960
atgtccatga aggaggtgga cgagcagatg ctggccatcc agagcaagaa cagcagctac 1020
ttcgtggagt ggatcccaa caacgtgaag gtggcctgt gtgacatccc gccccgcggc 1080
ctcaagatgt cctccacctt catcggaac agcacggcca tccaggagct gttcaagcgc 1140
atctccgagc agttcacggc catgttccgg cgcaaggcct tctgcaactg gtacacgggc 1200
gagggcattg acgagatgga gttcacggag gccgagagca acatgaacga cctgggtccc 1260
gagtaccagc agtaccagga cgccacggcc gaggaagagg gcgagatgta cgaagacgac 1320
gaggaggagt cggaggccca gggcccaag tgaactgct cgcagctgga gtgagaggca 1380
ggtggcggcc ggggccgaag ccagcagtggt ctaaaccccc ggagccatct tgctgccgac 1440
accctgcttt ccccatcgcc ctagggtccc cttgccgccc tctgcagta tttatggcct 1500
cgtcctcccc cacctaggcc acgtgtgagc tgctcctgtc tctgtcttat tgcagctcca 1560
ggcctgacgt tttacggttt tggtttttac tggtttgtgt ttatatattc ggggatactt 1620
aataaatcta ttgctgtcag ataccctt 1648

```

<210> 198

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006086

<400> 198

```

tttttactgg tttgtgttta tatttttcggg gataacttaat aaatctattg ctgtcagata 60

```

<210> 199

<211> 3074

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006096

<400> 199

```

aacaacacct gcctggctcc cagctggtgc tgaagctcgt cagttcacca tccgccctcg 60
gcttccgcgg ggcgctgggc cgccagcctc ggcaccgtcc ttctctttct ccctcgcggt 120
aggcaggtga cagcagggac atgtctcggg agatgcagga tgtagacctc gctgaggtga 180
agcctttggt ggagaaaggg gagaccatca ccggcctcct gcaagagttt gatgtccagg 240
agcaggacat cgagacttta catggetctg ttacggtcac gctgtgtggg actcccaagg 300
gaaaccggcc gtcatcctc acctaccatg acatcgcat gaaccacaaa acctgctaca 360
acccctcttt caactacgag gacatgcagg agatcaccca gcactttgcc gtctgccacg 420
tggacgcccc tggccagcag gacggcgag cctccttccc cgcagggtac atgtaccctt 480
ccatggatca gctggctgaa atgcttctct gagtccttca acagtttggg ctgaaaagca 540
ttattggcat gggaacagga gcaggcgctt acatcctaac tcgatttgct ctaaacaacc 600
ctgagatggt ggagggcctt gtccttatca acgtgaaccc ttgtgcggaa ggctggatgg 660
actgggccgc ctccaagatc tcaggatgga cccaagctct gccggacatg gtggtgtccc 720
acctttttgg gaaggaagaa atgcagagta acgtggaagt ggtccacacc taccgccagc 780
acattgtgaa tgacatgaac cccggcaacc tgcacctgtt catcaatgcc tacaacagcc 840

```

```

ggcgcgacct ggagattgag cgaccaatgc cgggaaccca cacagtcacc ctgcagtgcc 900
ctgctctgtt ggtggttggg gacagctcgc ctgcagtgga tgccgtggtg gagtgcact 960
caaaattgga cccaacaaag accactctcc tcaagatggc ggactgtggc ggcctccgc 1020
agatctccca gccggccaag ctgctgagg ccttcaagta cttcgtgcag ggcattggat 1080
acatgccctc ggctagcatg accgcctga tgcggtccc cagagcctct ggttccagcg 1140
tactttctct ggatggcacc cgcagccgct ccacaccag cgagggcacc cgaagccgct 1200
ccacaccag cgagggcacc cgcagccgct cgacaccag cgagggggcc cactggaca 1260
tcaccccaa ctcggtgct gctgggaaca ggcggggcc caagtccat gaggtctct 1320
gctaggcggc ctgcccagct gccgccccg gactctgatc tctgtagtgg cccctctct 1380
cccgccccc tttgcccc tgcctgccat actgcgccta actcggtatt aatccaaagc 1440
ttattttgta agagtggagct ctggtggaga caaatgaggt ctattacgtg ggtgccctct 1500
ccaaaggcgg ggtggcgtg gaccaaagga aggaagcaag catctccgca tcgcatcctc 1560
ttccattaac cagtggcgg ttgccactct cctccctcc ctacagagaca ccaaactgcc 1620
aaaaacaaga cgcgtagcag cacacacttc acaaagccaa gcctaggccg ccctgagcat 1680
cctggttcaa acgggtgcct ggtcagaagg ccagccgcc acttccggt tctctttaa 1740
ctgaggagaa gctgatccag tttccgaaa caaatcctt ttctcatttg gggagggggg 1800
taatagtac atgcaggcac ctcttttaa caggcaaaac aggaagggg aaaaggtggg 1860
attcatgtcg aggtagagg catttggaa acaaatcta cgtagttaac ttgaagaaac 1920
cgatttttaa agttggtgca tctagaaagc ttgaatgca gaagcaaaca agcttgattt 1980
ttctagcatc ctctaatgt gcagcaaaag caggcgacaa aatctcctgg ctttacagac 2040
aaaaatattt cagcaaacgt tgggcatcat ggtttttgaa ggcttttagt ctgctttctg 2100
cctctcctcc acagcccaa cctccacccc ctgatacatg agccagtgat tattcttggt 2160
cagggagaag atcatttaga tttgttttgc attccttaga atggagggca acattccaca 2220
gtgcccgtgg ctgtgatgag tgtccttgca ggggcggag taggagcact ggggtggggg 2280
tggaattggg gttactcgat gtaagggatt ccttgttgtt gtgttgagat ccagtgcagt 2340
tgtgatttct gtggatccca gcttggttcc aggaattttg tgtgattggc ttaaateccag 2400
ttttcaatct tcgacagctg ggctggaacg tgaactcagt agctgaacct gtctgacctg 2460
gtcacgttct tggatcctca gaactccttg ctcttgtcgg ggtgggggtg ggaactcacg 2520
tggggagcgg tggctgagaa aatgtaagga ttctggaata catattccat gggactttcc 2580
ttccctctcc tgcctcctct tttcctgctc cctaaccttt cgccgaatgg ggcagacca 2640
ctgacgtttc tgggcggcca gtgcggctgc caggttcctg tactactgcc ttgtactttt 2700
cattttggct caccgtggat tttctcatag gaagtttggc cagagtgaat tgaatattgt 2760
aagtcagcca ctgggacctg aggatttctg gaccccgca gttgggagga ggaagtagtc 2820
cagccttcca ggtggcgtga gaggcaatga ctcgttacct gccgccatc accttgagg 2880
ccttccttgg ccttgagtag aaaagtcggg gatcggggca agagaggctg agtacggatg 2940
ggaaactatt gtgcacaagt ctttcagag gagtttctta atgagatatt tgtattttatt 3000
tcagaccaaa taaatttgta actttgcagc ggaaaaaaaa aaaaaaaa aaaaaaaa 3060
aaaaaaaaa aaaa 3074

```

<210> 200

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006096

<400> 200

gagtacggat gggaaactat tgtgcacaag tctttccaga ggagtttctt aatgagatat 60

<210> 201

<211> 2148

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006115

<400> 201

```

gcttcagggt acagctcccc cgcagccaga agcggggcct gcagcgctc agcacggctc 60
cgggacaccc caccgccttc ccaggcgtga cctgtcaaca gcaacttcgc ggtgtggtga 120

```

```

actctctgag gaaaaacat tttgattatt actctcagac gtgcgtggca acaagtgact 180
gagacctaga aatccaagcg ttggagggtcc tgaggccagc ctaagtcgct tcaaaatgga 240
acgaaggcgt ttgtggggtt ccattcagag ccgatacatc agcatgagtg tgtggacaag 300
cccacggaga cttgtggagc tggcagggca gagcctgctg aaggatgagg ccctggccat 360
tgccgccctg gagttgtgc ccaggagct cttcccgcca ctcttcattg cagcctttga 420
cggaagacac agccagaccc tgaaggcaat ggtgcaggcc tggcccttca cctgcctccc 480
tctgggagtg ctgatgaagg gacaacatct tcacctggag accttcaaa cctgtgcttga 540
tggacttgat gtgctccttg cccaggaggt tcgcccagc aggtggaaac ttcaagtgtc 600
ggatttacgg aagaactctc atcaggactt ctggactgta tggctcggaa acagggccag 660
tctgtactca tttccagagc cagaagcagc tcagcccatg acaaagaagc gaaaagtaga 720
tggtttgagc acagaggcag agcagccctt cattccagta gaggtgctcg tagacctgtt 780
cctcaaggaa ggtgcctgtg atgaattgtt ctcctacctc attgagaaag tgaagcgaaa 840
gaaaaatgta ctacgcctgt gctgtaagaa gctgaagatt tttgcaatgc ccatgcagga 900
tatcaagatg atcctgaaaa tgggtgcagct ggactctatt gaagatttgg aagtgacttg 960
tacctggaag ctacccacct tggcgaaatt ttctccttac ctgggcccaga tgattaatct 1020
gcctagactc ctcctctccc acatccatgc atcttcttac atttccccgg agaaggaaga 1080
gcagtatatc gccaggttca cctctcagtt cctcagtcctg cagtgcctgc aggtctctta 1140
tgtggactct ttatttttcc ttagaggccg cctggatcag ttgctcaggc acgtgatgaa 1200
ccccttgga accctctcaa taactaactg ccggctttcg gaaggggatg tgatgcatct 1260
gtccagagt cccagcgta gtcagctaag tgtcctgagt ctaagtgggg tcatgctgac 1320
cgatgtaagt cccgagcccc tccaagctct gctggagaga gcctctgcca cctccagga 1380
cctggtcttt gatgagtgtg ggatcacgga tgatcagctc cttgccctcc tgccttccct 1440
gagccactgc tcccagctta caaccttaag cttctacggg aattccatct ccatatctgc 1500
cttgacagat ctccctgcagc acctcatcgg gctgagcaat ctgaccacg tgetgtatcc 1560
tgtcccccctg gagagttatg aggacatcca tggtagcctc cacctggaga ggcttgcta 1620
tctgcatgcc aggtcaggg agttgctgtg tgagttgggg cggcccagca tggctcggct 1680
tagtgccaac cctgtcctc actgtgggga cagaaccttc tatgaccgg agcccatcct 1740
gtgccctgt ttcatgccta actagctggg tgcacatatc aaatgcttca ttctgcatac 1800
ttggacacta aagccaggat gtgcatgcat cttgaagcaa caaagcagcc acagtttcag 1860
acaaatgttc agtgtgagtg aggaaaacat gttcagtgag gaaaaaacat tcagacaaat 1920
gttcagtgag gaaaaaaagg ggaagttggg gataggcaga tgttgacttg aggagttaat 1980
gtgatctttg gggagatata tcttatagag ttagaaatag aatctgaatt tctaaaggga 2040
gattctggct tgggaagtac atgtaggagt taatccctgt gtagactgtt gtaaagaaac 2100
tgttgaaaat aaagagaagc aatgtgaagc aaaaaaaaa aaaaaaaa 2148

```

<210> 202

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006115

<400> 202

```

tggggagata catcttatag agttagaat agaacttgaa tttctaaagg gagattctgg 60

```

<210> 203

<211> 1051

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006332

<400> 203

```

ggaccgccgc ctggttaaag gcgcttattt cccaggcagc cgctgcagtc gccacacctt 60
tgcccctgct gcgatgacct tgtcgccact tctgctgttc ctgccaccgc tgetgctgct 120
gttgagcgtc cccacggcgg cgggtgcaggc gtccctctct caagcgtag acttcttttg 180
gaatggggcca ccagtttaact acaagacagg caatctatac ctgcgggggc cctgaagaa 240
gtccaatgca ccgcttgta atgtgacctt ctactatgaa gcaactgtgc gtggctgcgc 300

```


PCT/US2006/004280

```

agccttcctg atccgggagc tcttcccaac atggctgttg gtcattggaga tcttcaatgt 360
cacgctgggt ccctacggaa acgcacagga acaaaatgtc agtggcaggt gggagttcaa 420
gtgccagcat ggagaagagg agtgcaaatt caacaagggt gaggcctgcg tgttgatga 480
acttgacatg gagctagcct tcttgaccat tgtctgcatg gaagagtttg aggacatgga 540
gagaagtctg ccactatgcc tgcagctcta cgccccaggg ctgtcgccag acactatcat 600
ggagtgtgca atgggggacc gcggcatgca gctcatgcac gccaacgccc agcggacaga 660
tgctctccag ccaccacacg agtatgtgcc ctgggtcacc gtcaatggga aacctttgga 720
agatcagacc cagctcctta cccttgcttg ccagttgtac cagggcaaga agccggatgt 780
ctgccccttc tcaaccagct ccctcaggag tgtttgcttc aagtgtatggc cgggtgagctg 840
cggagagctc atggaaggcg agtgggaacc cggctgcctg cctttttttc tgatccagac 900
cctcggcacc tgctacttac caactggaaa attttatgca tcccatgaag ccagatata 960
caaaattcca ccccatgatc aagaatcctg ctccactaag aatgggtgcta aagtaaaact 1020
agtttaataa gcaaaaaaaa aaaaaaaaaa a 1051

```

<210> 204

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006332

<400> 204

```

aaattccacc cctagatcaa gaatcctgct ccactaagaa tgggtgctaaa gtaaaactag 60

```

<210> 205

<211> 1714

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006417

<400> 205

```

ggggcatttt gtgcctgcct agctatccag acagagcagc taccctcagc tctagctgat 60
actacagaca gtacaacaga tcaagaagta tggcagtgac aactcgtttg acacgggttg 120
acgaaaagat cctgcaaaat cattttggag ggaagcggct tagccttctc tataagggtta 180
gtgtccatgg attccgtaat ggagttttgc ttgacagatg ttgtaatcaa gggcctactc 240
taacagtgat ttatagttaa gatcatatta ttggagcata tgcggaagag agttaccagg 300
aaggaaagta tgcttccatc atcctttttg cacttcaaga tactaaaatt tcagaatgga 360
aactaggact atgtacacca gaaacactgt tttgttgtga tgttacaaaa tataactccc 420
caactaattt ccagatagat ggaagaaata gaaaagtgat tatggactta aagacaatgg 480
aaaatcttgg acttgctcaa aattgtacta tctctattca ggattatgaa gtttttcgat 540
gcgaagattc actggatgaa agaaagataa aaggggtcat tgagctcagg aagagcttac 600
tgtctgcctt gagaacttat gaaccatatt gatccctggt tcaacaaata cgaattctgc 660
tgctgggtcc aattggagct gggaagtcca gctttttcaa ctcagtgagg tctgttttcc 720
aagggcatgt aaogcatcag gctttggtgg gcaactaata aactgggata tctgagaagt 780
ataggacata ctctattaga gacgggaaag atggcaaata cctgccgttt attctgtgtg 840
actcactggg gctgagtgag aaagaaggcg gcctgtgcag ggatgacata ttctatatct 900
tgaacggtaa cattcgtgat agataccagt ttaatcccat ggaatcaatc aaattaaatc 960
atcatgacta cattgattcc ccatacgtga aggacagaat tcattgtgtg gcattttgat 1020
ttgattgccg ctctattcaa tacttctcct ctcatgatgt agtaaagatc aaaagaattc 1080
gaagggaggt ggtaaacgct ggtgtggtac atgtggccttt gctcactcat gtggatagca 1140
tggaattgat taaaaaagg gaccttatag aaatagagag atgtgagcct gtgaggtcca 1200
agctagagga agtccaaaga aaacttggat ttgctctttc tgacatctcg gtggttagca 1260
attattcctc tgagtgggag ctggaccctg taaaggatgt tctaattctt tctgctctga 1320
gacgaatgct atgggctgca gatgacttct tagaggattt gccttttgag caaataggga 1380
atctaaggga ggaaattatc aactgtgcac aaggaaaaaa atagatatgt gaaaggttca 1440
cgtaaatctc ctcacatcac agaagattaa aattcagaaa ggagaaaaca cagaccaaag 1500
agaagtatct aagaccaaag ggatgtgttt tattaatgtc taggatgaag aaatgcatag 1560
aacattgtag tacttgtaaa taactagaaa taacatgatt tagtcataat tgtgaaaaat 1620

```

agtaataatt tttcttggat ttatgttctg tatctgtgaa aaaataaatt tcttataaaa 1680
ctcggaaaaa aaaaaaaaaa aaaaaaaaaa aaaa 1714

<210> 206
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006417

<400> 206
atgacatatt ctatatcttg aacggtaaca ttcgtgatag ataccagttt aatcccatgg 60

<210> 207
<211> 3791
<212> DNA
<213> Homo sapiens

<300>
<308> NM_006461

<400> 207
acagacggcg ggtgaacatg gcgtcctcga cttgggtctga gacgtgatag gcctgccttc 60
tggttgaaga tgtggcgagt gaaaaaactg agcctcagcc tgtcgccttc gccccagacg 120
ggaaaaccat ctatgagaac tcctctccgt gaacttaccg tgcagcccgg tgccctcacc 180
acctctggaa aaagatcccc cgcttgcctc tcgtgaccc catcactgtg caagctgggg 240
ctgcaggaag gcagcaacaa ctctctcca gtggattttg taaataacaa gaggacagac 300
ttatcttcag aacatttcag tcattcctca aagtggctag aaacttgtca gcatgaatca 360
gatgagcagc ctctagatcc aattccccaa attagctcta ctccataaac gtctgaggaa 420
gcagtagacc cactgggcaa ttatatgggt aaaaaccatcg tccttgtagc atctccactg 480
gggcagcaac aagacatgat atttgaggcc cgtttagata ccatggcaga gacaaacagc 540
atatctttaa atggaccttt gagaacagac gatctgggtg gagaggaggt ggccacctgc 600
atgggagaca ggttttcaga agttgtctgt gtatctgaga aacctatctt tcaggaatct 660
ccgtcccatc tcttagagga gtctccacca aatccctggt ctgaacaact acattgctcc 720
aaggaaagcc tgagcagtag aactgaggct gtgcgtgagg acttagtagc ttctgaaagt 780
aacgccttct tgcccttcctc tgttctctgt ctttccctct caactgcctt ggcagcagat 840
ttccgtgtca atcatgtgga cccagaggag gaaattgtag agcatggagc tatggaggaa 900
agagaaatga ggtttccac acatcctaag gagtctgaaa cagaagatca agcacttgtc 960
tcaagtgtgg aagatattct gtccacatgc ctgacaccaa atctagtaga aatggaatcc 1020
caagaagctc caggcccagc agtagaagat gttggttaga ttcttggtc tgatacagag 1080
tcttgatgt cccactggc ctggctggaa aaagggtgaa atacctcgt catgctggaa 1140
aatctccgcc aaagcttctc ccttccctcg atgcttcggg atgctgcaat tggcactacc 1200
cctttctcta ctgtctcggg ggggaacttg tttactcctt cagcaccaca ggaaaagagt 1260
acaaacacat cccagacagg cctggttggc accaagcaca gtacttctga gacagagcag 1320
ctctgtgtg gccggcctcc agatctgact gcctgtctc gacatgactt ggaagataac 1380
ctgctgagct ctctgtcat tgtggagttt ctctcccgcc agcttcggga ctggaagagc 1440
cagctggctg tcctcaccg agaaaccag gacagtagca cacagactga cacatctcac 1500
agtgggataa ctaataaact tcagcatctt aaggagagcc atgagatggg acaggcccta 1560
cagcaggcca gaaatgtcat gcaatcatgg gtgcttatct ctaaagagct gatatctctg 1620
cttcacctat ccctgttgca tttagaagaa gataagacta ctgtgaatca ggagtctcgg 1680
cgtgcagaaa cattggtctg ttgctgtttt gatttgctga agaaattgag ggcaaagctc 1740
cagagcctca aagcagaaag ggaggaggca aggcacagag aggaaatggc tctcagaggc 1800
aaggatgcgg cagagatagt gttggaggct ttctgtgcac acgccagcca gcgcacagc 1860
cagctggaac aggacctagc atccatgcgg gaattcagag gccttctgaa ggatgccag 1920
acccaactgg tagggcttca tgccaagcaa gaagagctgg ttcagcagac agtgagctct 1980
acttctacct tgcaacaaga ctggagggtcc atgcaactgg attatacaac atggacagct 2040
ttgtgagtc ggtcccgaca actcacagag aaactcacag tcaagagcca gcaagccctg 2100
caggaaactg atgtggcaat tgaggaaaag caggagggtt ctagggtgct ggaacaagtc 2160
tctgccagat tagaggagt caaaggccaa acagaacaac tggagttgga aaacattcgt 2220
ctagcaacag atctccgggc tcagttgcag attctggcca acatggacag ccagctaaaa 2280

```

gagctacaga gtcagcatac ccattgtgcc caggacctgg ctatgaagga tgagttactc 2340
tgccagctta cccagagcaa tgaggagcag gctgctcaat gcgtaaagga agagatggca 2400
ctaaaacaca tgcaggcaga actgcagcag caacaagctg tcctggccaa agaggtgagg 2460
gacctgaaag agaccttga gtttgcagac caggagaatc aggttgctca cctggagctg 2520
ggtcaggttg agtgtcaatt gaaaaccaca ctggaagtgc tccgggagcg cagcttgcag 2580
tgtgagaacc tcaaggacac tgtagagaac ctaacggcta aactggccag caccatagca 2640
gataaccagg agcaagatct ggagaaaaca cggcagtact ctcaaaagct agggctgctg 2700
actgagcaac tacagagcct gactctcttt ctacagacaa aactaaagga gaagactgaa 2760
caagagaccc ttctgctgag tacagcctgt cctcccaccc aggaacaccc tctgcctaact 2820
gacaggacct tcctgggaag catcttgaca gcagtggcag atgaagagcc agaacaact 2880
cctgtgccct tgcttggaag tgacaagagt gctttcaccg gagtagcatc aatggtttcc 2940
cttcagcccg cagagacccc aggcattggg gagagcctgg cagaaatgag tattatgact 3000
actgagcttc agagtctttg ttccctgcta caagagtcta aagaagaagc catcaggact 3060
ctgcagcgaa aaatttgtga gctgcaagct aggttgccag cccaggaaga acagcatcag 3120
gaagtccaga aggcaaaaaga agcagacata gagaagctga accaggcctt gtgcttgccg 3180
tacaagaatg aaaaggagct ccagggaagt atacagcaga atgagaagat cctagaacag 3240
atagacaaga gtggcgagct cataagcctt agagaggagg tgaccacact taccgctca 3300
cttcggcggtg cggagacaga gacaaaagtg ctccaggagg cctggcaggc cagctggact 3360
ccaactgccg gcctatggcc accaattgga tccaggagaa agtgtggctc tctcaggagg 3420
tggaacaaact gagagtgatg ttccctggaga tgaaaaatga gaaggaaaac tcctgatcaa 3480
gttccagagc ccatagaaat atcctagagg agaaccttcg gcgctctgac aaggagttag 3540
aaaaactaga tgacattgtt cagcatatct ataagacctt gctctctatt ccagaggtgg 3600
tgaggggatg caaagaacta cagggattgc tggaatttct gagctaagaa actgaaagcc 3660
agaatttgtt tcacctcttt ttacctgcaa tacccttcta cccaataacc aagaccaact 3720
ggcatagagc caactgagat aaatgctatt taaataaagt gtatttaatg aaaaaaaaaa 3780
aaaaaaaaaa a 3791

```

<210> 208
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006461

<400> 208
 ctgacaagga gttagaaaaa ctagatgaca ttgttcagca tatttataag accctgctct 60

<210> 209
 <211> 2856
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006516

<400> 209
 tagtcgcggg tccccgagtg agcacgccag ggagcaggag accaaacgac gggggtcgga 60
 gtcagagtcg cagtgggagt ccccgaccg gagcacgagc ctgagcggga gacgcgcgct 120
 cgcacgcccg tcgccaccgg cgtaccgggc gcagccagag ccaccagcgc agcgctgcca 180
 ttggagcccag cagcaagaag ctgacgggtc gcctcatgct ggctgtggga ggagcagtg 240
 ttggctccct gcagtttgcc tacaacactg gactcatcaa tgccccccag aaggtgatcg 300
 aggagttcta caaccagaca tgggtccacc gctatgggga gagcatcctg cccaccacgc 360
 tcaccacgct ctggctccctc tcagtggcca tcttttctgt tgggggcatg attggctcct 420
 tctctgtggg ccttttctgt aaccgctttg gccggcgga ttcaatgctg atgatgaacc 480
 tgctggcctt cgtgtccgcc gtgctcatgg gcttctcgaa actgggcaag tcctttgaga 540
 tgctgatcct gggcgccttc atcatcggtg tgtaactgcg cctgaccaca ggcttcgtgc 600
 ccatgtatgt gggatgaagt tcacccacag cctttcgtgg ggccctgggc accctgcacc 660
 agctgggcat cgtcgtcggc atcctcatcg ccagggtgtt cggcctggac tccatcatgg 720
 gcaacaagga cctgtggccc ctgctgctga gcatcatctt catcccgcc ctgctgcagt 780
 gcatcgtgct gcccttctgc cccgagagtc cccgcttctt gctcatcaac cgcaacgagg 840

```

agaaccgggc caagagtgtg ctaaagaagc tgcgcgggac agctgacgtg acccatgacc 900
tgcaggagat gaaggaagag agtcggcaga tgatgcggga gaagaaggtc accatcctgg 960
agctgttccg ctcccccgcc taccgccagc ccatacctcat cgctgtggtg ctgcagctgt 1020
cccagcagct gtctggcatc aacgctgtct tctattactc cagcagcatc ttcgagaagg 1080
cgggggtgca gcagcctgtg tatgccacca ttggctccgg tatcgtcaac acggccttca 1140
ctgtcgtgtc gctgtttgtg gtggagcgag caggccggcg gaccctgcac ctcataggcc 1200
tcgttggcat ggcggttgtt gccatactca tgaccatcgc gctagcactg ctggagcagc 1260
taccctggat gtcttatctg agcatcgtgg ccatacttgg ctttgtggcc ttctttgaag 1320
tgggtcctgg ccccatccca tgggtcatcg tggctgaact cttcagccag ggtccacgtc 1380
cagctgccat tgccgttgca ggcttctcca actggacctc aaatttcatt gtgggcattg 1440
gcttccagta tgtggagcaa ctgtgtggtc cctacgtctt catcatcttc actgtgctcc 1500
tggttctgtt cttcatcttc acctacttca aagttcctga gactaaaggc cggaccttcg 1560
atgagatgcg ttccggcttc cggcaggggg gagccagcca aagtataag acacccgagg 1620
agctgttcca tcccttgggg gctgattccc aagtgtgagt cgcctcagat caccagcccg 1680
gcctgtcccc agcagcccta aggatctctc aggagcacag gcagctggat gagacttcca 1740
aacctgacag atgtcagccg agccgggcct ggggctcctt tctccagcca gcaatgatgt 1800
ccagaagaat attcaggact taacggctcc aggattttaa caaaagcaag actgttgctc 1860
aaatctatct agacaagcaa caggttttat aattttttta ttactgattt tgttattttt 1920
atatcagcct gagtctcctg tgcccacatc ccaggcttca ccctgaatgg ttccatgcct 1980
gagggtggag actaagccct gtcgagacac ttgccttctt caccagcta atctgtaggg 2040
ctggacctat gtcttaagga cacactaatc gaactatgaa ctacaaagct tctatccag 2100
gagggtggcta tggccaccgg ttctgtctgc ctggatctcc ccactctagg ggtcaggctc 2160
cattaggatt tgccccttcc catctcttcc taccacaacca ctcaaattaa tctttcttta 2220
cctgagacca gttgggagca ctggagtgcg gggaggagag gggaagggcc agtctgggct 2280
gccgggttct agtctccttt gcactgaggg ccacactatt accatgagaa gagggcctgt 2340
gggagcctgc aaactcactg ctcaagaaga catggagact cctgccctgt tgtgtataga 2400
tgcaagatat ttatatatat ttttggttgt caatatataa tacagacact aagttatagt 2460
atatctggac aagccaactt gtaaatacac cacctcactc ctgttactta cctaaacaga 2520
tataaatggc tggtttttag aaacatggtt ttgaaatgct tgtggattga gggtaggagg 2580
tttggatggg agtgagacag aagtaagtgg ggttgcaacc actgcaacgg cttagacttc 2640
gactcaggat ccagtcctct acacgtacct ctcatcagtg tcctcttgcg caaaaatctg 2700
tttgatccct gttaccacaga gaatatatac attctttatc ttgacattca aggcatttct 2760
atcacatatt tgatagttgg tgttcaaaaa aacactagtt ttgtgccagc cgtgatgctc 2820
aggcttgaat tcgcattatt ttgaatgtga agggaa 2856

```

<210> 210

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006516

<400> 210

```

aaacagatat aaatggctgg tttttagaaa catggttttg aaatgcttgt ggattgaggg 60

```

<210> 211

<211> 576

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006607

<400> 211

```

atggctactc tgatctacgt tgataaggaa attggagaac caggcaccgg tgtggctgcc 60
aaggatgtgc tgaagctgga gtctagacct tcaatcaaag cattagatgg gatctctcaa 120
gttttaacac caggttttgg caaaacatac gatgctccat cagccttacc taaagctacc 180
agaaaggctt tgggcaactg caacagagct acagaaaagt cagtaaagac caatggaccc 240
agaaaacaaa aacagccaag cttttctgcc aaaaagatga ccgagaagac tgtaaaaaca 300
aaaagtctcg ttctgcctc agatgacgcc tatccagaaa tagaaaaatt ctttcccttc 360

```

```

aatcttctag actttgagag ttttgacctg cctgaagagc gccagattgc acacctcccc 420
ttgagtggag tgcctctcat gatccttgat gaggaggag agcttgaaaa gctgtttcag 480
ctgggcccc cttcacctgt gaaaatgccc tctccaccat gggaatgcaa tctgtttgca 540
gtctccttca agcattctgt cgaccctgga tgttga 576

```

```

<210> 212
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_006607

```

```

<400> 212
cgccatcca gaaatagaaa aattctttcc cttcaatctt ctagactttg agagttttga 60

```

```

<210> 213
<211> 2058
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_006820

```

```

<400> 213
gcacgaggaa gccacagatc tcttaagaac tttctgtctc caaaccgtag ctgctcgata 60
aatcagacag aacagttaat cctcaattta agcctgatct aacctctaga aacagatata 120
gaacaatgga agtgacaaca agattgacat ggaatgatga aaatcatctg cgcaactgct 180
tggaatggtt tctttgagtc ttctctataa gtctagtgtt catggaggta gcattgaaga 240
tatggttgaa agatgcagcc gtcagggatg tactataaca atggottaca ttgattacaa 300
tatgattgta gcctttatgc ttggaaatta tattaattta cgtgaaagtt ctacagagcc 360
aaatgattcc ctatggtttt cacttcaaaa gaaaaatgac accactgaaa tagaaacttt 420
actcttaa atacagaccaa aaattattga tgagcaactg gtgtgtcgtt tatcgaaaac 480
ggatattttt attatatgtc gagataataa aatttatcta gataaaatga taacaagaaa 540
cttgaaaacta aggttttatg gccaccgtca gtatttggaa tgtgaagttt ttcgagttga 600
aggaattaag gataacctag acgacataaa gaggataaatt aaagccagag agcacagaaa 660
taggcttcta gcagacatca gagactatag gccctatgca gacttggttt cagaaattcg 720
tattcttttg ttgggtccag ttgggtccag aaagtccagt tttttcaatt cagtcaagtc 780
tatttttcat ggccatgtga ctggccaagc cgtagtgggg tctgatacca ccagcataac 840
cgagcgggat aggatataat ctgttaaaga tggaaaaaat ggaaaatctc tgccatttat 900
gttggtgtgac actatggggc tagatggggc agaaggagca ggactgtgca tggatgacat 960
tccccacatc ttaaaagggt gtatgccaga cagatatcag tttaattccc gtaaaccaat 1020
tacacctgag cattctactt ttatcacctc tccatctctg aaggacagga ttcactgtgt 1080
ggcttatgtc ttagacatca actctattga caatctctac tctaaaatgt tggcaaaagt 1140
gaagcaaggt cacaagaag tattaaactg tggtagtagc tatgtggcct tgcttactaa 1200
agtggatgat tgcagtgagg ttcttcaaga caacttttta aacatgagta gatctatgac 1260
ttctcaaagc cgggtcatga atgtccataa aatgctaggc attcctattt ccaatatattt 1320
gatggttga aattatgctt cagatttga actggacccc atgaaggata ttctcatcct 1380
ctctgcactg aggcagatgc tgcgggctgc agatgatttt ttagaagatt tgcctcttga 1440
ggaaactggt gcaattgaga gagcgttaca gccctgcatt tgagataagt tgccttgatt 1500
ctgacatttg gccagcctg tactgggtgt ccgcaatgag agtcaatctc tattgacagc 1560
ctgcttcaga ttttgctttt gttcgttttg ccttctgtcc ttggaacagt catatctcaa 1620
gttcaaaggc caaaacctga gaagcgggtg gctaagatag gtccactgac aaaccacccc 1680
tccatatttc cgtaccattt acaattcagt ttctgtgaca tctttttaaa cactggagg 1740
aaaaatgaga tattctctaa ttattcttc tataacactc tatatagagc tatgtgagta 1800
ctaatacat tgaataatag ttataaaatt attgtataga catctgcttc ttaaacagat 1860
tgtgagttct ttgagaaaca gcgtggattt tacttatctg tgtattcaca gagcttagca 1920
cagtgcctgg taatgagcaa gcatacttgc cattactttt ccttcccact ctctccaaca 1980
tcacattcac tttaaatttt tctgtatata gaaaggaaaa ctagcctggg caacatgatg 2040
aaaccccatc tccactgc 2058

```

Patent Application No. 2006/084272

<210> 214
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006820

<400> 214
 tgagttcttt gagaaacagc gtggatttta cttatctgtg tattcacaga gcttagcaca 60

<210> 215
 <211> 2825
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006845

<400> 215
 gcgaaattga gggttcttgg tattgcgcgt ttctcttcct tgctgactct ccgaatggcc 60
 atggactcgt cgcttcaggc ccgcctgttt cccggctctg ctatcaagat ccaacgcagt 120
 aatggttttaa ttcacagtgc caatgtaagg actgtgaact tggagaaatc ctgtgtttca 180
 gtggaatggg cagaaggagg tgccacaaag ggcaaaagaga ttgattttga tgatgtggct 240
 gcaataaaac cagaactctt acagcttctt cccttacatc cgaaggacaa tctgcccttg 300
 caggaaaatg taacaatcca gaaacaaaaa cggagatccg tcaactccaa aattcctgct 360
 ccaaaagaaa gtcttcgaag ccgctccact cgcattgtcca ctgtctcaga gcttcgcac 420
 acggctcagg agaattgacat ggaggtggag ctgcctgcag ctgcaaaactc ccgcaagcag 480
 ttttcagttc ctctgcccc cactaggcct tcctgcccctg cagtggctga aataccattg 540
 aggatggatc gcgaggagat ggaagagcaa gtccattcca tccgtggcag ctcttctgca 600
 aaccctgtga actcagttcg gaggaatca tgtcttgtga aggaagtga aaaaatgaag 660
 aacaagcgag aagagaagaa ggcccagaa tctgaaatga gaatgaagag agctcaggag 720
 tatgacagta gttttccaaa ctgggaattt gcccgaaatg ttaaagaatt tcgggctact 780
 ttggaatgtc atccacttac tatgactgat cctatcgaag agcacagaat atgtgtctgt 840
 gttaggaaaac gccactgaa taagcaagaa ttggccaaga aagaaattga tgtgatttcc 900
 attcctagca agtgtctcct cttggtacat gaacccaagt tgaaagtga cttaacaaag 960
 tatctggaga accaagcatt ctgctttgac ttgcatcttg atgaaacagc ttcgaaatgaa 1020
 gttgtctaca ggttcacagc aaggccactg gtacagacaa tctttgaagg tggaaaagca 1080
 acttgttttg catatggcca gacagggaag ggcaagacac atactatggg cggagacctc 1140
 tctgggaaag ccagaaatgc atccaaaggg atctatgcca tggcctcccg ggacgtcttc 1200
 ctctgaaaga atcaaccctg ctaccggaag ttgggcctgg aagtctatgt gacattcttc 1260
 gagatctaca atgggaagct gtttgacctg ctcaacaaga aggccaaagct gcgcgtgctg 1320
 gaggacggca agcaacaggt gcaagtgtgt gggctgcagg agcatctggt taactctgct 1380
 gatgatgtca tcaagatgct cgacatgggc agcgcctgca gaacctctgg gcagacattt 1440
 gccaaactcca attcctcccg ctcccacgcg tgcttccaaa ttattcttcg agctaaaggg 1500
 agaatgcatg gcaagtcttc tttggtagat ctggcaggga atgagcgagg cgcagacact 1560
 tccagtgtg accggcagac ccgcatggag ggccgagaaa tcaacaagag tctcttagcc 1620
 ctgaaggagt gcatcagggc cctgggacag aacaaggctc acaccccgtt ccgtgagagc 1680
 aagctgacac aggtgctgag ggactccttc attggggaga actctaggac ttgcatgatt 1740
 gccacgatct caccaggcat aagctcctgt gaataactt taaacaccct gagatatgca 1800
 gacagggtca agagctgag cccccacagt gggccacagt gagagcagtt gattcaaatg 1860
 gaaacagaag agatggaagc ctgcttaac gggtcgctga ttccaggcaa ttatocaaag 1920
 gaagaggagg aactgtcttc ccagatgtcc agctttaacg aagccatgac tcagatcagg 1980
 gagctggagg agaaggctat ggaagagctc aaggagatca tacagcaagg accagactgg 2040
 cttgagctct ctgagatgac cgagcagcca gactatgacc tggagacctt tgtgaacaaa 2100
 gcggaatctg ctctggccca gcaagccaag catttctcag ccctgcgaga tgtcatcaag 2160
 gccttacgcc tggccatgca gctggaagag caggctagca gacaaataag cagcaagaaa 2220
 cggcccccagt gacgactgca aataaaaaatc tggttggttt gacacccagc ctcttccctg 2280
 cggctcccca gagaactttg ggtacctggt gggcttaggc aggtctgtg ctgggacagg 2340
 ttctggtaaa tgccaagtat gggggcatct gggccaggga cagctgggga gggggtcaga 2400

```

gtgacatggg acactccttt tctgttcctc agttgtcgcc ctcacgagag gaaggagctc 2460
ttagttaccc ttttgtgttg cccttctttc catcaagggg aatgtttctc gcatagagct 2520
ttctccgcag catcctgcct gcgtggactg gctgctaatt gagagctccc tggggttgct 2580
ctggctctgg ggagagagac ggagccttta gtacagctat ctgctggctc taaaccttct 2640
acgcctttgg gccgagcact gaatgtcttg tactttaaaa aaatgtttct gagacctctt 2700
tctactttac tgtctcccta gagtcctaga ggatccctac tgttttctgt tttatgtgtt 2760
tatacattgt atgtaacaat aaagagaaaa aataaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2820
aaaaa 2825

```

```

<210> 216
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_006845

```

```

<400> 216
aaatgtttct gagacctctt tctactttac tgtctcccta gagtcctaga ggatccctac 60

```

```

<210> 217
<211> 823
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_007019

```

```

<400> 217
aaacgcgggc gggcgggccc gcagtcctgc agttgcagtc gtgtttctccg agttcctgtc 60
tctctgccaa cgccgcccgg atggcttccc aaaaccgcga cccagccgcc actagcgtcg 120
ccgccgcccg taaaggagct gagccgagcg gggcgccgcg ccgggggtccg gtgggcaaaa 180
ggctacagca ggagctgatg accctcatga tgtctggcga taaagggatt tctgccttcc 240
ctgaatcaga caaccttttc aaatgggtag ggaccatcca tggagcagct ggaacagtat 300
atgaagacct gaggtataag ctctcgctag agttccccag tggctaccct tacaatgcgc 360
ccacagtga gttcctcacg cctgtctatc accccaacgt ggacaccag ggtaacatat 420
gcctggacat cctgaaggaa aagtggctcg ccctgtatga tgtcaggacc attctgctct 480
ccatccagag ccttctagga gaaccaaca ttgatagtc cttgaacaca catgctgccg 540
agctctggaa aaaccccaca gcttttaaga agtacctgca agaaacctac tcaaagcagg 600
tcaccagcca ggagccctga cccaggctgc ccagcctgtc cttgtgtcgt ctttttaatt 660
tttccttaga tggctctgtc tttttgtgat ttctgtatag gactctttat cttgagctgt 720
ggtatttttg ttttgttttt gtcttttaaa ttaagcctcg gttgagccct tgtatattaa 780
ataaatgcat ttttgtcctt ttttagacaa aaaaaaaaaa aaa 823

```

```

<210> 218
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_007019

```

```

<400> 218
tggaataaac ccacagcttt taagaagtac ctgcaagaaa cctactcaaa gcaggtcacc 60

```

```

<210> 219
<211> 2831
<212> DNA
<213> Homo sapiens

```

```

<300>

```

<308> NM_007183

<400> 219

```

gaattccgga caggacgtga agatagttgg gtttggaggc ggccgccagg cccaggcccc 60
gtggacctgc cgccatgcag gacggtaact tectgtgtc ggccctgcag cctgaggccg 120
gcgtgtgctc cctggcgctg ccctctgacc tgcagctgga ccgcccgggc gccgaggggc 180
cggaggccga gcggctgcgg gcagcccgcg tccaggagca ggtccgcgcc gcctctttgc 240
agctgggaca gcagccgcgg cacaacgggg ccgctgagcc cgagcctgag gccgagactg 300
ccagaggcac atccaggggg cagtaccaca cctgcaggc tggcttcagc tctcgctctc 360
agggcctgag tggggacaag acctcgggct tccggcccat cgccaagccg gcctacagcc 420
cagcctcctg gtctctccgc tccgcctgg atctgagct cagtccgagg ctgagttcag 480
cccacaatgg gggcagcgcc tttggggccg ctgggtacgg ggggtgccag cccacccctc 540
ccatgccac caggcccggtg tcttccatg agcgcggtg ggttgggagc cgggccgact 600
atgacacact ctccctgcgc tcgctgcgcg tggggcccg gggcctggac gaccgctaca 660
gcctggtgtc tgagcagctg gagcccgcg ccacctccac ctacagggcc tttgcgtacg 720
agcgccaggc cagctccagc tccagccggg caggggggct ggactggccc gaggccactg 780
aggtttcccc gagccggacc atccgtgcc ctgccgtgc gacctgcag cgattccaga 840
gcagccacgg gagccgggg gtaggcgggg cagtgcggg ggccgtcctg gagccagtgg 900
ctcagcgccc atctgtgcgc agcctcagcc tcagcctggc tgactcgggc cacctgccgg 960
acgtgcatgg gttcaacagc tacggtagcc accgaacct gcagagactc agcagcggtt 1020
ttgatgacat tgacctgccc tcagcagtc agtacctcat ggcttcagac cccaacctgc 1080
aggtgctggg agcggcctac atccagcaca agtgctacag cgatgcagcc gccagaagc 1140
agggccgcag ccttcaggcc gtgcctaggc tgggtgaagc cttcaaccac gccaacagg 1200
aagtgcagcg ccagtcgaca ggtgccatgc gcaacctcat ctacgacaac gctgacaaca 1260
agctggccct ggtggaggag aacgggatct tcgagctgct gcggacactg cgggagcagg 1320
atgatgagct tcgcaaaaat gtcacaggga tctgtggaa cctttcatcc agcgaccacc 1380
tgaaggaccg cctggccaga gacacgctgg agcagctcac ggacctggtg ttgagcccc 1440
tgtcgggggc tgggggtccc cccctcatcc agcagaacgc ctcgaggcg gagatcttct 1500
acaacgccac cggcttctc aggaacctca gctcagcctc tcaggccact cgccagaaga 1560
tgccggagtg ccacgggctg gtggacgccc tggtcacctc tatcaaccac gccctggacg 1620
cgggcaaatg ccaggaaca agcgtggaga acgcggtgtg cgtcctgcgg aacctgtcct 1680
accgctcta cgacgagatg ccgccgtccg cgtgcagcg gctggagggt cgcggccgca 1740
gggacctggc gggggcgccg ccgggagagg tcgtgggctg cttcacgccg cagagccggc 1800
ggctgcgcga gctgccctc gccgccgat cgtcacctt cgcggagggt tccaaggacc 1860
ccaaggccct cgagtggctg tggagcccc agatcgtggg gctgtacaac cggctgctgc 1920
agcgtgcga gctcaaccg cacacgacgg aggcggccgc cggggcgctg cagaacatca 1980
cggcaggcga ccgcagggtg gcgggggtgc tgagccgct ggccctggag caggagcgta 2040
ttctgaacct cctgtagag cgtgtcagga ccgcccagca ccaccagctg cgctcactga 2100
ctggcctcat ccgaaacctg tctcggaacg ctaggaaaca ggacgagatg tccacgaagg 2160
tggtgagcca cctgatcgag aagctgccag gcagcgtggg tgagaagtgc cccccagccg 2220
aggtgctggt caacatcata gctgtgctca acaacctggt ggtggccagc cccatcgctg 2280
cccagacact gctgtatttt gacggactcc gaaagctcat ctcatcaag aagaagcggg 2340
acagccccga cagtgagaag tctcccggg cagcatccag cctcctggcc aacctgtggc 2400
agtacaacaa gctccaccgt gactttcggg cgaagggtta tcggaaggag gacttctctg 2460
gcccataagg gaagccttct ggaggagaag gtgacgtggc ccagcgtcca agggacagac 2520
tcagctccag gctgcttgcc agcccagcct ggaggagaag gctaatacag gaggggcccc 2580
tcgctggggc cctgtgtgc atctttgagg gtccctgggc accaggaggg gcagggtctt 2640
atagctgggg acttggttc cgcagggcag ggggtggggc agggctcaag gctgctctgg 2700
tgtatggggg ggtgaccag tcacattggc agagggtggg gttggctgtg gcctggcagt 2760
atcttgggat agccagcact gggaataaag atggccatga acagtcacaa aaaaaaaaaa 2820
aaaaggaatt c 2831

```

<210> 220

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007183

<400> 220

ctggcagtat cttgggatag ccagcactgg gaataaagat ggccatgaac agtcacaaaa 60

<210> 221

<211> 2815

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007267

<400> 221

aggaagcgga ggaaggtgaa gtaggaccga attcctgtgc cgaagaggcc tgcagtggga 60
gagcaggatg ggggctccgg aggtggcgcc caggctctga gctaccctag gtctgcagac 120
tagcgggcat tggccagaga catggcccag ccactggcct tcatcctcga tgtccctgag 180
accccagggg accagggcca gggcccccag ccctatgatg aaagcgaagt gcacgactcc 240
ttccagcagc tcatccagga gcagagccag tgcacggccc aggaggggct ggagctgcag 300
cagagagagc gggaggtgac aggaagtagc cagcagacac tctggcggcc cgagggcacc 360
cagagcagcg ccacactccg catcctggcc agcatgccc gccgcacat tggcgcgagc 420
cgaggtgcca tcatctccca gtactacaac cgcacgggtg agcttcgggtg caggagcagc 480
cggccctctg tcgggaactt tgtccgctcc gcctggccca gcctccgctt gtacgacctg 540
gagctggacc ccacggccct ggaggaggag gagaagcaga gcctcctggt gaaggagtcc 600
cagagcctgg cagtggcaca gcgggaccac atgcttcgctg ggatgccctt aagcctggct 660
gagaacgcga gcctgcgaga gaagagcagg accccgaggg ggaagtggag gggccagccg 720
ggcagcgcg ggtctgtctc ctgtctgtggc cggctcagat atgcctgcgt gctggccttg 780
cacagcctgg ccctggcgct gctctccgcc ctgcaggccc tgatgccgtg gcgtacagc 840
ctgaagcgca tcgggggcca gttcggctcc agcgtgctct cctacttctt ctttctcaag 900
acctgtctgg ctttcaatgc cctcctgtct ctgtctgtgg tggccttcat catgggccct 960
caggtgcctt tcccacccgc cctgccgggc cctgccccgc tctgcacagg cctggagctc 1020
ctcacaggcg cgggttgctt caccacacac gtcatgtact acggccacta cagtaacgcc 1080
acgtgaacc agcctgtgtg cagccccctg gatggcagcc agtgcacacc cagggtgggt 1140
ggcctgccc acaacatgccc cctggcctac ctctccactg tgggcgtgag cttctttatc 1200
acctgcatca ccctggtgta cagcatggct cactctttcg gggagagcta ccgggtgggc 1260
agcacctctg gcatccacgc catcacctgc ttctgtctct gggactacaa ggtgacgcag 1320
aagcgggccc cccgcctcca gcaggacaat attcgacccc ggctgaagga gctgctggcc 1380
gagtggcagc tgcggcacag ccccaggagc gtgtgcggga ggctgcggca ggcggctgtg 1440
ctggggcttg tgtggctgct gtgtctgggg accgcgctgg gctgcgcctg gcccgctcac 1500
gtcttctcgg agttcatgat ccagagtcca gaggtgctg gccaggaggc tgtgctgctg 1560
gtcctgccc tgggtggttg cctcctcaac ctggggggccc cctacctgtg ccgtgtcctg 1620
gccgccctgg agccgcatga ctccccgcta ctggaggtgt acgtggccat ctgcaggaa 1680
ctcatcctca agctggccat cctggggaca ctgtgctacc actggctggg ccgcagggtg 1740
ggcgtcctgc agggccagtg ctgggaggat tttgtggggc aggagctgta ccggttctctg 1800
gtgatggact tcgtcctcat gttgctggac acgctttttg gggaaactgt gtggaggatt 1860
atctccgaga agaagctgaa gaggaggcgg aagccggagt ttgacattgc ccggaatgtc 1920
ctggagctga tttatgggca gactctgacc tggctggggg tgcctttctc gccctcctc 1980
cccgccgtgc agatcatcaa gctgctgctc gtcttctatg tcaagaagac cagccttctg 2040
gccaaactgcc aggcgcgcgc ccggccctgg ctggcctcac acatgagcac cgtcttctc 2100
acgtgctctt gcttccccgc cttcctgggc gcgctgtct tctctgcta cgcgctctg 2160
cagtggaagc cctcgagcac ctgcggcccc ttccggaccc tggacacccat gtacgaggcc 2220
ggcagggtgt ggggtgcgcca cctggaggcg gcaggcccca gggctcctct gctgccctgg 2280
gtgcaccggg acctgatgga aaacaccttc tttgtcttcc tgggtgtcagc cctgctgctg 2340
gccgtgatct acctcaacat ccaggtggtg cggggccagc gcaaggtcat ctgcctgctc 2400
aaggagcaga tcagcaatga gggtaggagc aaaatcttct taatcaacaa gcttactcc 2460
atctacgaga ggaaggagag ggaggagag agcagggttg ggacaaccga ggaggtgcg 2520
gcacccctcg cctgtctcac agatgaacag gatgcctagg gggacggcga tgggcctcac 2580
gggcccgcgc agcacctga gaccacactg ttgcctccca gtgacctgc tgggacacca 2640
ggacaaggaa gacagtctcg cctctcgaaa gccgcagctg cgcctaggct ggagctggaa 2700
gggtgggtga atccggcttg ggcaccccca atgaactctg ccctgcctgg gactctatct 2760
attctgatta aaggggtttt gcaaatggga aaaaaaaaa aaaaaaaaa aaaaa 2815

<210> 222

<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> NM_007267

<400> 222
ggtgaggaca aaatcttctt aatcaacaag cttcactcca tctacgagag gaaggagagg 60

<210> 223
<211> 1893
<212> DNA
<213> Homo sapiens

<300>
<308> NM_007274

<400> 223
atttaccgcc gcgcggagag tgagggccca agtcgcgcct gctccgccac ttaggcgcgc 60
ccagacgctt ccctcggggc tgccaccggg tcgggcgcgc ctgccgcggc tagcgggcct 120
tccccgcacc gggcgcggcc aaccgccacc gaaccttctg gaagcggcgg ctgcctgggc 180
ccccacgccg ccagaatcgt acgcccgcgc gagctctctg cagccttggc gccctgggag 240
gcggggctcg ggggtggggc ggcgcggggg cggggtcggc gcgggggagg cgcgttcgat 300
tcgcccccg gcgcgaggcc cgcctcacc agccccatcg ctccacctct gccctcccc 360
tttatggcgc ggcccgggct cattcattcc gcgcggggc tgccagacac ctgcgcctt 420
ctgcagccgc ccgcgcgcatc cgccgcgcga gccccagca tgtcggggcc agacgtcgag 480
acgcccgtcc ccattccgat ctgccggatc atgcggccag atgatgccaa cgtggccggc 540
aatgtccacg gggggacccat cctgaagatg atcgaggagg caggcgccat catcagcacc 600
cggcattgca acagccagaa cggggagcgc tgtgtggccg ccctggctcg tgtcgagcgc 660
accgacttcc tgtctcccat gtgcatcggg gaggtggcgc atgtcagcgc ggagatcacc 720
tacacctcca agcactctgt ggaggtgcag gtcaacgtga tgtccgaaaa catcctcaca 780
ggtgccaaaa agctgaccaa taaggccacc ctgtggtatg tgcccctgtc gctgaagaat 840
gtggacaagg tctcgcaggt gcctcctgtt gtgtattccc ggcaggagca ggaggaggag 900
ggccggaagc ggtatgaagc ccagaagctg gagcgcatgg agaccaagtg gaggaacggg 960
gacatcgctc agccagtcct caaccagag ccgaacactg tcagctacag ccagtccagc 1020
ttgatccacc tgggtggggc ttcagactgc accctgcacg gctttgtgca cggagggtgtg 1080
accatgaagc tcatggatga ggtcgccggg atcgtggctg cacgccactg caagaccaac 1140
atcgtcacag cttccgtgga cgccattaat ttcatgaca agatcagaaa aggctgcgtc 1200
atcaccatct cgggacgcat gaccttcacg agcaataagt ccatggagat cgagggtgtg 1260
gtggacgccc accctgttgt ggacagctct cagaagcgct accgggcccgc cagtgccttc 1320
ttcacctacg tgtcgcgtgag ccaggaaggc aggtcgctgc ctgtgccccg gctggtgccc 1380
gagaccgagg acgagaagaa gcgctttgag gaaggcaaag ggcggtacct gcagatgaag 1440
gcgaagcgac agggccacgc ggagcctcag ccctagactc cctcctcctg ccactgggtg 1500
ctcagtagtc catggcaacg ggcccagtgt ccagtcactt agaagttccc cccttgggca 1560
aaaaccacaat tcacattgag agctgggtgt gtctgaagtt ttcgtatcac agtggttaacc 1620
tgtactctct cctgcaaaacc tacacaccaa agctttatct atatcattcc agtatcaatg 1680
ctacacagtg ttgtcccagc cgccgggagg cgttgggcag aaaccctcgg gaatgcttcc 1740
gagcacgctg tagggtagg gaagaaccca gcaccactaa taaagctgct gcttggctgg 1800
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1860
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa 1893

<210> 224
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_007274

<400> 224
acctacacac caaagcttta tttatatcat tccagtatca atgctacaca gtgttgtccc 60

<210> 225
 <211> 4157
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007315

<400> 225
 agcggggcgg ggcgccagcg ctgccttttc tctgcccggg tagtttcgct ttcttgccga 60
 gagtctgcgg aggggctcgg ctgcaccggg gggatcgcgc ctggcagacc ccagaccgag 120
 cagagcgagc ccagcgcgct cgggagaggc tgcaccgcgc cgccccgcc tagcccttcc 180
 ggatcctgcg cgcagaaaag tttcatttgc tgtatgccat cctcgagagc tgtctagggt 240
 aacgttcgca ctctgtgtat ataacctcga cagtcttggc acctaacgtg ctgtgcgtag 300
 ctgtcctttt ggttgaatcc ccaggccctt gttggggcac aaggtggcag gatgtctcag 360
 tggtaacgaac ttcagcagct tgactcaaaa ttcttgagc aggttcacca gctttatgat 420
 gacagttttc ccatggaaat cagacagtac ctggcacagt ggtagaaaa gcaagactgg 480
 gagcacgctg ccaatgatgt ttcatttggc accatccgtt ttcatgacct cctgtcacag 540
 ctggatgatc aatatagtcg cttttctttg gagaataact tcttgctaca gcataacata 600
 aggaaaagca agcgtaatct tcaggataat tttcaggaag acctaatcca gatgtctatg 660
 atcatttaca gctgtctgaa ggaagaaagg aaaattctgg aaaacgcccc gagatttaat 720
 caggctcagt cggggaatat tcagagcaca gtgatgttag acaaacagaa agagcttgac 780
 agtaaaagta gaaatgtgaa ggacaagggt atgtgtatag agcatgaaat caagagcctg 840
 gaagatttac aagatgaaata tgacttcaaa tgcaaaaacct tgcagaacag agaacacgag 900
 accaatggtg tggcaaaagag tgatcagaaa caagaacagc tgttactcaa gaagatgtat 960
 ttaatgcttg acaataagag aaaggaagta gttcacaaaa taatagagtt gctgaatgtc 1020
 actgaactta cccagaatgc cctgattaat gatgaactag tggagtggaa gcggagacag 1080
 cagagcgctt gtattggggg gccgccccat gcttgcttgg atcagctgca gaactgggtc 1140
 actatagttg cggagagtct gcagcaagtt cggcagcagc ttaaaaagtt ggaggaattg 1200
 gaacagaaat acacctacga acatgacctt atcacaaaaa acaacaagt gttatgggac 1260
 cgcaccttca gtcttttcca gcagctcatt cagagctcgt ttgtggtgga aagacagccc 1320
 tgcatgccaa cgcacctca gaggccgctg gtcttgaaga caggggtcca gttcactgtg 1380
 aagttagac tggttggtgaa attgcaagag ctgaattata atttgaaagt caaagtctta 1440
 tttgataaag atgtgaatga gagaaatata gtaaaaggat ttaggaagtt caacattttg 1500
 ggcacgcaca caaaagtgat gaacatggag gagtccacca atggcagctt gccggtgaa 1560
 tttcggcacc tgcaattgaa agaacagaaa aatgctggca ccagaacgaa tgagggctct 1620
 ctcatcgtaa ctgaagagct tcactccctt agttttgaaa cccaattgtg ccagcctggt 1680
 ttgtaattg acctcgagac gacctctctg ccggttggg tgatctccaa cgtcagccag 1740
 ctcccagcgg gttgggcttc catcctttgg tacaacatgc tgggtggcga acccaggaat 1800
 ctgtccttct tctgactcc accatgtgca cgatgggctc agctttcaga agtgcgtgag 1860
 tggcagtttt cttctgtcac caaaagaggt ctcaatgtgg accagctgaa catggtggga 1920
 gagaagcttc ttggtcctaa cgcagcccc gatggtctca ttccgtggac gagggtttgt 1980
 aaggaaaata taaatgataa aaattttccc ttctggcttt ggattgaaag catcctagaa 2040
 ctcatataaa aacacctgct cctctcttgg aatgatgggt gcatcatggg cttcatcagc 2100
 aaggagcgag agcgtgccct gttgaaggac cagcagccgg ggaccttctt gctgcggttc 2160
 agtgagagct cccgggaagg ggccatcaca ttcacatggg tggagcggtc ccagaacgga 2220
 ggcgaacctg acttccatgc ggttgaacct tacacgaaga aagaactttc tgcgtgtact 2280
 ttccctgaca tcattcgcaa ttacaaagtc atggctgctg agaattttcc tgagaatccc 2340
 ctgaagtatc tgtatccaaa tattgacaaa gacctgcctt ttggaaagta ttactccagg 2400
 ccaaaggaag caccagagcc aatggaactt gatggcccta aaggaactgg atatatcaag 2460
 actgatttga tttctgtgtc tgaagttcac cctctagac ttcagaccac agacaacctg 2520
 ctcccatgt ctctgagga gtttgacgag gtgtctcgga tagtgggctc tgtagaattc 2580
 gacagtatga tgaacacagt atagagcatg aatttttttc atcttctctg gcgacagttt 2640
 tccttctcat ctgtgattcc ctctgctac tctgttctct cactcctgt gtttctaggg 2700
 aaatgaaaga aaggccagca aattcgctgc aacctgttga tagcaagtga atttttctct 2760
 aactcagaaa catcagttac tctgaagggc atcatgcatc ttactgaagg taaaattgaa 2820
 aggcattctc tgaagagtgg gtttcacaag tgaaaaacat ccagatacac ccaaagtatc 2880
 aggacgagaa tgagggctct ttgggaaagg agaagttaag caacatctag caaatgttat 2940
 gcataaagtc agtgcccaac tgttataggt tgttgataaa atcagtggtt atttaggga 3000
 ctgcttgacg taggaacggt aaatttctgt gggagaattc ttacatgttt tctttgcttt 3060

```

aagtgttaact ggcagttttc cattgggtta cctgtgaaat agttcaaagc caagttttata 3120
tacaattata tcagtcctct ttcaaaggta gccatcatgg atctggtagg gggaaaatgt 3180
gtatttttatt acatctttca cattggctat ttaaagacaa agacaaattc tgtttcttga 3240
gaagagaata ttagctttac tgtttgttat ggcttaatga cactagctaa tatcaataga 3300
aggatgtaca ttcccaaatt cacaagttgt gtttgatatc caaagctgaa tacattctgc 3360
tttcatcttg gtcacatata attattttta cagttctccc aaggaggtta ggctattcac 3420
aaccactcat tcaaaagtgt aaattaacca tagatgtaga taaactcaga aatttaattc 3480
atgtttctta aatgggctac tttgtccttt ttgttattag ggtgggtattt agtctattag 3540
ccacaaaatt gggaaaggag tagaaaaagc agtaactgac aacttgaata atacaccaga 3600
gataatatga gaatcagatc atttcaaac tcatctccta tgtaactgca ttgagaactg 3660
catatgtttc gctgatatat gtgtttttca catttgcgaa tggttccatt ctctctcctg 3720
tactttttcc agacactttt ttgagtggat gatgtttcgt gaagtatact gtatttttac 3780
ctttttcctt ccttatcact gacacaaaaa gtagattaag agatgggttt gacaagggtc 3840
ttccctttta catactgctg tctatgtggc tgtatcttgt tttccacta ctgctaccac 3900
aactatatta tcatgcaaat gctgtattct tctttgggtg agataaagat ttcttgagtt 3960
ttgttttaaa attaaagcta aagtatctgt attgcattaa atataatat cacacagtgc 4020
ttccgtggc actgcataca atctgaggcc tctctctca gtttttata agatggcgag 4080
aacctaagtt tcagttgatt ttacaattga aatgactaaa aaacaaagaa gacaacatta 4140
aaacaatatt gtttcta 4157

```

<210> 226

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007315

<400> 226

```

atcagatcat ttcaaaactc atttcctatg taactgcatt gagaactgca tatgtttcgc 60

```

<210> 227

<211> 1696

<212> DNA

<213> Homo sapiens

<300>

<308> NM_009587

<400> 227

```

caaaggactt cctagtgggt gtgaaaggca ggggtggcca cagaggcggc ggagagatgg 60
ccttcagcgg ttcccagcct ccctacctga gtccagctgt ccccttttct gggactattc 120
aaggaggtct ccaggacgga cttcagatca ctgtcaatgg gaccgttctc agctccagtg 180
gaaccaggtt tgcgtgaac tttcagactg gtttcagtgg aaatgacatt gccttccact 240
tcaaccctcg gtttgaagat ggagggtagc tgggtgtgcaa cagcaggcag aacggaagct 300
ggggggccga ggagaggaag acacacatgc tttccagaa ggggatgcc tttgacctct 360
gcttctcgtt gcagagctca gatttcaagg tgatggtgaa cgggatcctc ttcgtgcagt 420
acttccaccg cgtgcccttc caccgtgtgg acaccatctc cgtcaatggc tctgtgcagc 480
tgtctacat cagcttccag aacccccgca cagtccctgt tcagcctgcc ttctccacgg 540
tgccgttctc ccagcctgtc tgtttcccac ccaggcccag ggggcccaga caaaaacctc 600
ccggcgtgtg gcctgccaac ccggctccca ttaccagac agtcatccac acagtgcaga 660
gcgccctcg acagatgttc tctactcccg ccatcccacc tatgatgtac cccacccccg 720
cctatccgat gcctttcatc accaccatc tgggagggtc gtaccatcc aagtccatcc 780
tctgtcagg cactgtcctg ccagtgctc agaggttcca catcaacctg tgcctctggga 840
accacatgc cttccacctg aacccccgtt ttgatgagaa tgctgtggtc cgcaacaccc 900
agatcgacaa ctctggggg tctgaggagc gaagtctgcc ccgaaaaatg cccttcgtcc 960
gtggccagag cttctcagtg tggatcttgt gtgaagctca ctgcctcaag gtggccgtgg 1020
atggtcagca cctgtttgaa tactaccatc gcctgaggaa cctgccacc atcaacagac 1080
tggaaagtgg ggcggacatc cagctgaccc atgtgcagac ataggcggct tcctggccct 1140
ggggccgggg gctggggtgt ggggcagtct gggctcctc atcatccca cttcccaggc 1200
ccagccttcc caaccctgcc tgggatctgg gctttaatgc agaggccatg tccttgtctg 1260

```

```

gtcctgcttc tggctacagc caccctggaa cggagaaggg agctgacggg gattgccttc 1320
ctcagccgca gcagcacctg gggctccagc tgctggaatc ctaccatccc aggaggcagg 1380
cacagccagg gagaggggag gagtgggcag tgaagatgaa gcccctgct cagtccctc 1440
ccatccccca cgcagctcca cccagtccc aagccaccag ctgtctgctc ctggtgggag 1500
gtggcctcct cagccctcc tctctgacct ttaacctcac tctcaccttg caccgtgcac 1560
caacccttca cccctcctgg aaagcaggcc tgatggcttc ccactggcct ccaccacctg 1620
accagagtgt tctcttcaga ggactggctc ctttcccagt gtccttaaaa taaagaaatg 1680
aaaatgcttg ttggca 1696

```

<210> 228
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_009587

<400> 228
 cagaggactg gctcctttcc cagtgtcctt aaaataaaga aatgaaaatg cttgttgcca 60

<210> 229
 <211> 6552
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_012291

<400> 229
 atgaggagct tcaaaagagt caactttggg actctgctaa gcagccagaa ggaggctgaa 60
 gagttgctgc cgcagctgaa ggagttcctg tccaaccctc cagctggttt tcccagcagc 120
 cgatctgatg ctgagaggag acaagcttgt gatgccatcc tgagggcttg caaccagcag 180
 ctgactgcta agctagcttg ccctaggcat ctggggagcc tgctggagct ggcagagctg 240
 gcctgtgatg gctacttagt gtctacccca cagcgtcctc cctctacct ggaacgaatt 300
 ctctttgtct tactgcggaa tgctgctgca caaggaagcc cagaggccac actccgcctt 360
 gctcagcccc tccatgcctg cttggtgcag tgctctcgcg aggtgtctcc ccaggactat 420
 gaggccgtgg ctgggggcag cttttctctg ctttgggaagg gggcagaagc cctgttgga 480
 cggcgagctg catttgcagc tcggctgaag gccttgagct tccatgtact cttggaggat 540
 gaaagtaccc cttgtgaggt tccctacttt gcttctccaa cagcctgtcg agcggtagct 600
 gcccatcagc tatttgatgc cagtggccat ggtctaaatg aagcagatgc tgatttccta 660
 gatgacctgc tctccaggca cgtgatcaga gccttggtgg gtgagagagg gagctcttct 720
 gggcttcttt ctccccagag ggccctctgc ctcttgagc tcaccttgga aactgcccgt 780
 cgcttttgct ggagccgccca ccatgacaaa gccatcagcg cagtggagaa ggctcacagt 840
 tacctaagga acaccaatct agcccctagc cttcagctat gtcagctggg ggttaagctg 900
 ctgcaggttg gggaggaagg acctcaggca gtggccaagc ttctgatcaa ggcacagct 960
 gtcctgagca agagtatgga ggcacatca ccccacttc gggcattgta tgagagctgc 1020
 cagttctctc tttcaggcct ggaacgaggg accaagaggg gctatagact tgatgccatt 1080

ctgagcctct ttgcttttct tggagggtac tgctctcttc tgcagcagct gcgggatgat 1140
 ggtgtgtatg ggggctcctc caagcaacag cagtcttttc ttcatgatga ctttcaggga 1200
 cttcacctct acactgtggt ggtttatgac tttgcccagg gctgtcagat agttgatattg 1260
 gctgacctga cccaactagt ggacagttgt aaatctaccg ttgtctggat gctggaggcc 1320
 ttagagggcc tgctggggcca agagctgacg gaccacatgg ggatgacgc ttcttacacc 1380
 agtaatttgg cctacagctt ctatagtcac aagctctatg ccgaggcctg tgccatctct 1440
 gagccgctct gtcagcacct gggtttggtg aagccaggca cttatcccga ggtgcctcct 1500
 gagaagtgtc acaggtgctt ccggctacaa gttagagagt tgaagaaact gggtaaacag 1560
 gccaggggct gcaagatggt gatthttgtg ctggcagccc tgcaaccctg tagccctgaa 1620
 cacatggctg agccagtcac tttctgggtt cgggtcaaga tggatgcggc cagggtcgga 1680
 gacaaggagc tacagctaaa gactctgcga gacctacaag cgtgctggga cccggagacc 1740
 ctggccctcc tgctgagga ggagctgcag gctacaagg cgtgctggga cgacactgga 1800
 caggaacgct tcaacatcat ctgtgacctc ctggagctga gcccgagga gacaccagcc 1860

ggggcctggg	cacgagccac	ccacctggta	gaactggctc	aggtgctctg	ctaccacgac	1920
tttacgcagc	agaccaactg	ctctgctctg	gatgotatcc	gggaagccct	gcagcttctg	1980
gactctgtga	ggcctgaggc	ccaggccaga	gatcagcttc	tggacgataa	agcacaggcc	2040
ttgtctggc	tttacatctg	tactctggaa	gcaaaaatac	aggaaggtat	cgagcgggat	2100
cggagagccc	aggcccctgg	taacttggag	gaatttgaag	tcaatgacct	gaactatgaa	2160
gataaactcc	aggaagatcg	tttccataac	agtaacattg	ccttcaacct	ggctgcagat	2220
gctgctcagt	ccaaatgcct	ggaccaagcc	ctggccctgt	ggaaggagct	gcttacaagg	2280
gggcaggccc	cagctgtacg	gtgtctccag	cagacagcag	cctcaactga	gatccatgca	2340
gccctctacc	agctggtggc	aaagcccatg	caggctctgg	aggtcctcct	gctgctacgg	2400
attgtctctg	agagactgaa	ggaccactcg	aaggcagctg	gctcctcctg	ccacatcacc	2460
cagctcctcc	tgaccctcgg	ctgtcccagc	tatgccagct	tacacctgga	agaggcagca	2520
tcgagcctga	agcatctcga	tcagactact	gacacatacc	tgctcctttc	cctgacctgt	2580
gatctgcttc	gaagtcaact	ctactggact	caccagaagg	tgaccaaggg	tgtctctctg	2640
ctgctgtctg	tgtctcggga	tccctgccctc	cagaagtcct	ccaaggcttg	gtacttgctg	2700
cgtgtccagg	tccctgcagct	ggtggcagct	tacotttagcc	tcccgtcaaa	caacctctca	2760
cactccctgt	gggagcagct	ctgtgcccaa	ggctggcaga	cacctgagat	agctctcata	2820
gactcccata	agctcctccg	aagcatcatc	ctcctgtcga	tgggcagtga	cattctctca	2880
actcagaaag	cagctgtgga	gacatcgttt	ttggactatg	gtgaaaatct	ggtacaaaaa	2940
tggcagggtt	tttcagaggt	gctgagctgc	tcagagaagc	tggtctgcca	cctgggccgc	3000
ctgggtagtg	tgagtgaagc	caaggccttt	tgtttggagg	ccctaaaact	tacaacaaag	3060
ctgcagatac	cacgccagtg	tgccctgttc	ctgggtgctga	agggcgagct	ggagctggcc	3120
cgcaatgaca	ttgatctctg	tcagtcggac	ctgcagcagg	ttctgttctt	gcttgagtct	3180
tgacacagat	ttggtggggt	gactcagcac	ctggactctg	tgaagaaggt	ccacctgcag	3240
aagggggaag	agcaggccca	ggtccctgtg	cctccacagc	tcccagagga	ggagctcttc	3300
ctaagaggcc	ctgctctaga	gtggttgccc	actgtggcca	aggagcctgg	ccccatagca	3360
ccttctacaa	actcctcccc	agtcttgaaa	accaagcccc	agcccatacc	caacttctcg	3420
tcccattcac	ccacctgtga	ctgctcgttc	tgcgccagcc	ctgtcctcac	agcagtctgt	3480
ctgcctggg	tatttgtcac	ggcaggggtg	aggctggcca	tgggccacca	agcccagggt	3540
ctggatctgc	tgacggctgt	gctgaagggc	tgtcctgaag	ccgctgagcg	cctcacccaa	3600
gctctccaag	cttccctgaa	tcataaaaaca	ccccctcct	tggttccaag	cctcttggat	3660
gagatcttgg	ctcaagcata	cacactgttg	gcactggagg	gctgaacca	gccatcaaac	3720
gagagcctgc	agaagttctt	acagtcaggg	ctgaagtttg	tagcagcacg	gataccccac	3780
ctagagccct	ggcgagccag	cctgctcttg	atttggggcc	tcacaaaact	aggtggcctc	3840
agctgctgta	ctacccaact	ttttgcaagc	tccctggggct	ggcagccacc	attaataaaa	3900
agtgtccctg	gctcagagcc	ctctaagact	cagggccaaa	aacgttcttg	acgagggcgc	3960
caaaagttag	cctctgctcc	cctgcgcttc	aataatacct	ctcagaaaag	tctggaaggt	4020
agaggactgc	cctgcacacc	taaaccacca	gaccggatca	ggcaagctgg	ccctcatgtc	4080
cccttcacgg	tggttaggga	agtctgcctt	acagagagca	agcctgaagt	accccaggcc	4140
cccagggtac	aacagagagt	ccagacgcgc	ctcaagggtg	acttcaagtga	tgacagtgc	4200
ttggaagacc	ctgtctcagc	tgaggcctgg	ctggcagagg	agcctaagag	acggggcact	4260
gcttcccggg	gccgggggcg	agcaaggaag	ggcctgagcc	taaagacgga	tgccgtgggt	4320
gccccaggta	gtgcccttg	gaaccctggc	ctgaatggca	ggagccggag	ggccaagaag	4380
gtggcatcaa	gacattgtga	ggagcggcgt	ccccagaggg	ccagtgacca	ggccaggcct	4440
ggccctgaga	tcattgagac	catccctgag	gaagaactga	ctgacaactg	gagaaaaatg	4500
agctttgaga	tcctcagggg	ctctgacggg	gaagactcag	cctcaggttg	gaagactcca	4560
gctccggggc	ctgaggcagc	ttctggagaa	tgggagctgc	tgaggctgga	ttccagcaag	4620
aagaagctgc	ccagcccatg	cccagacaag	gagagtgaca	aggaccttgg	tcctcggttc	4680
cagctccctc	cagcccccgt	agccactggg	ctttctaccc	tggactccat	ctgtgactcc	4740
ctgagtgttg	ctttccgggg	cattagtcc	tgctcctccta	gtgggctcta	tgcccacctc	4800
tgccgcttcc	tgcccttggt	cctggggccac	cgggatcctt	atgccactgc	tttcccttgtc	4860
accaggtctg	tctccatcac	ctgtcgccac	cagctgtctc	cccacctcca	cagacagctc	4920
agcaaggccc	agaagcaccc	aggatcactt	gaaatagcag	accagctgca	ggggctgagc	4980
cttcaggaga	tgcttgagga	tgtcccccgt	gcccgcaccc	agcgctctt	ttccttcagg	5040
gctttggaat	ctggccactt	ccccagcct	gaaaaggaga	gtttccagga	gcgcctggct	5100
ctgatcccca	gtggggtgac	tgtgtgtgtg	ttggcccttg	ccacctcca	gcccggaaac	5160
gtgggcaaca	ccctcctgct	gaccgggctg	gaaaaggaca	gtccccaggt	cagtgtgcag	5220
attcccactg	gccagaacaa	gcttcatctg	cgttcagtc	tgaatgagtt	tgatgccatc	5280
cagaaggcac	agaaagagaa	cagcagctgt	actgacaagc	gagaatgggt	gacagggcgg	5340
ctggcaactg	accacaggat	ggaggttctc	atcgcttccc	tagagaagtc	tgtgctgggc	5400
tgttggaagg	ggctgctgct	gccgtccagt	gaggagcccg	gccctgccca	ggaggcctcc	5460

```

cgccctacagg agctgctaca ggactgtggc tggaaatata ctgaccgcac tctgctgaaa 5520
atcatgctca gtggtgccgg tgcctcacc cctcaggaca ttcaggccct ggcctacggg 5580
ctgtgcccaa cccagccaga gcgagcccag gagctcctga atgaggcagt aggacgtcta 5640
cagggcctga cagtaccaag caatagccac ctgtcttgg tcctagacaa ggacttgacag 5700
aagctgccgt gggaaagcat gcccagcctc caagcactgc ctgtcaccgc gctgccctcc 5760
ttccgcttcc tactcagcta tcccatcacc aaagagtatg gggcctcgcc agtgctgagt 5820
caaggggtgg atccacgaag taccttctat gtctgaacc ctcacaataa cctgtcaagc 5880
acagaggagc aatttcgagc caatttcagc agtgaagctg gctggagagg agtggttggg 5940
gaggtgccaa gacctgaaca ggtgcaggaa gccctgacaa agcatgattt gtatatctat 6000
gcagggcatg gggctggtgc ccgcttctt gatgggcagg ctgtcctgcg gctgagctgt 6060
cgggcagtgg cctgtctgtt tggctgtagc agtgcggccc tggctgtgca tggaaacctg 6120
gagggggctg gcctcgtgct caagtacatc atggctggtt gccccttgtt tctgggtaat 6180
ctctgggatg tgactgaccg cgacattgac cgctacacgg aagctctgct gcaaggctgg 6240
cttgagcag gcccaggggc ccccttctc tactatgtaa accaggcccg ccaagctccc 6300
cgactcaagt atcttattgg ggctgcacct atagcctatg gcttgctgt ctctctgcgg 6360
taaccccatg gagctgtctt attgatgcta gaagcctcat aactgttcta cctccaaggt 6420
tagatttaat ccttaggata actcttttaa agtgattttc cccagtgttt tatatgaaac 6480
atttcctttt gatttaacct cagtataata aagatacatc atttaaaccg tgaacaaaaa 6540
aaaaaaaaaa aa 6552

```

<210> 230

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_012291

<400> 230

```

agcctcataa ctgttctacc tccaagggtta gatttaatcc ttaggataac tcttttaaag 60

```

<210> 231

<211> 6317

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013261

<400> 231

```

tagtaagaca ggtgccttca gttcactctc agtaaggggc tgggtgcctg catgagtgtg 60
tgctctgtgt cactgtggat tggagttgaa aaagcttgac tggcgtcatt caggagctgg 120
atggcgtggg acatgtgcaa ccaggactct gactctgtat ggagtgcacat cgagtgtgct 180
gctctggttg gtgaagacca gcctctttgc ccagactctc ctgaacttga tctttctgaa 240
ctagatgtga acgacttgga tacagacagc ttctgtgggtg gactcaagtg gtgcagtgc 300
caatcagaaa taatatccaa tcagtacaac aatgagcctt caaacatatt tgagaagata 360
gatgaagaga atgaggcaaa cttgctagca gtccctcacag agacactaga cagtctccct 420
gtggatgaag acggattgcc ctcatctgat gcgctgacag atggagacgt gacctgac 480
aatgaggcta gtcttctctc catgcctgac ggcacccctc caccacagga ggcagaagag 540
ccgtctctac ttaagaagct cttactggca ccagccaaca ctgagctaag ttataatgaa 600
tgagtggtc tcagtagcca gaacctgca aatcacatc acaggatcag aacaaacctc 660
gcaattgtta agactgagaa ttcatggagc aataaagcga agagtatttg tcaacagcaa 720
aagccacaaa gacgtccctg ctggagcctt ctcaaatac tgaccacaaa cgatgacct 780
cctcacacca aaccacaga gaacagaaac agcagcagag acaaatgcac ctccaaaaa 840

```

```

aagtcacaca cacagtgcga gtcacaacac ttacaagcca aaccaacaac tttatctctt 900
cctctgacct cagagtccac aaatgacccc aagggttccc catttgagaa caagactatt 960
gaacgcacct taagtgtgga actctctgga actgcaggcc taactccacc caccactcct 1020
cctcataaag ccaaccaaga taacctttt agggcttctc caaagctgaa gtccctctgc 1080
aagactgtgg tgccaccacc atcaaagaag cccaggtaca gtgagtcttc tggtagacaa 1140

```

ggcaataact	ccaccaagaa	agggccggag	caatccgagt	tgtatgcaca	actcagcaag	1200
tcctcagtc	tcactggtg	acacgaggaa	aggaagacca	agcggccag	tctgcggtg	1260
tttgggtgacc	atgactattg	ccagtcaatt	aattccaaaa	cagaaatact	cattaatata	1320
tcacaggagc	tccaagactc	tagacaacta	gaaaataaag	atgtctctc	tgattggcag	1380
gggcagattt	gttcttccac	agattcagac	cagtgcctac	tgagagagac	tttggaggca	1440
agcaagcagg	tctctccttg	cagcacaaga	aaacagctcc	aagaccagga	aatccgagcc	1500
gagctgaaca	agcacttcgg	tcatcccagt	caagctgttt	ttgacgcaga	agcagacaag	1560
accggtgaac	tgaggggacag	tgatttcagt	aatgaacaat	tctccaaact	acctatgttt	1620

ataaattcag	gactagccat	ggatggcctg	tttcatgaca	gcgaagatga	aagtgataaa	1680
ctgagctacc	cttgggatgg	cacgcaatcc	tattcattgt	tcaatgtgtc	tccttcttgt	1740
tcttctttta	actctccatg	tagagattct	gtgtcaccac	ccaaatcctt	atcttctcaa	1800
agaccccaaa	ggatgcgctc	tcgttcaagg	tccttttctc	gacacaggtc	gtgttcccga	1860
tcaccatatt	ccagggtcaag	atcaagggtc	ccaggcgagta	gatcctcttc	aagatcctgc	1920
tattactatg	agtcaagcca	ctacagacac	cgacgcgacc	gaaattctcc	cttgtatgtg	1980
agatcacggt	caagatcgcc	ctacagccgt	cgccccagggt	atgacagcta	cgaggaatat	2040
cagcacgaga	ggctgaagag	ggaagaatat	cgaagagagt	atgagaagcg	agagctctgag	2100
agggccaagc	aaaggggagag	gcagaggcgag	aaggcaattg	aagagcgccg	tgtgatttat	2160
gtcggtaaaa	tcagacctga	cacaacacgg	acagaactga	gggaccgttt	tgaagttttt	2220
gggtgaaattg	aggagtgcac	agtaaactctg	cgggatgatg	gagacagcta	tggtttcatt	2280
acctaccgtt	atacctgtga	tgcttttctg	gctcttgaaa	atggatacac	tttgcgcagg	2340
tcaaacgaaa	ctgactttga	gctgtacttt	tgtggacgca	agcaattttt	caagtctaac	2400
tatgcagacc	tagattcaaa	ctcagatgac	tttgaccctg	ctccaccaaa	gagcaagtat	2460
gactctcttg	attttgatag	tttactgaaa	gaagctcaga	gaagcttgcg	caggtaacat	2520
gttccctagc	tgaggatgac	agagggatgg	cgaatacctc	atgggacagc	gcgtccttcc	2580
ctaaagacta	ttgcaagtca	tacttaggaa	tttctcctac	tttacactct	ctgtacaaaa	2640
acaaaacaaa	acaacaacaa	tacaacaaga	acaacaacaa	caataacaac	aatggtttac	2700
atgaacacag	ctgctgaaga	ggcaagagac	agaatgatat	ccagtaagca	catgtttatt	2760
catgggtgtc	agctttgctt	ttcctggagt	ctcttggtga	tggagtgtgc	gtgtgtgcat	2820
gtatgtgtgt	gtgtatgtat	gtgtgtgggt	tgtgtgcttg	gtttagggga	agtatgtgtg	2880
ggtagatctg	ggactgtggg	gcacctgacc	agaatgcgca	agggcaaacc	atttcaaattg	2940
gcagcagttc	catgaagaca	cgcttaaaac	ctagaacttc	aaaatgttcg	tattctattc	3000
aaaaggaaat	atatatatat	atatatatat	atatatatat	atatataaat	taaaaaggaa	3060
agaaaactaa	caaccaacca	accaaccaac	caaccacaaa	ccaccctaaa	atgacagccg	3120
ctgatgtctg	ggcatcagcc	tttgtactct	gtttttttta	gaaagtgcag	aatcaacttg	3180
aagcaagctt	tctctcataa	cgtaatgatt	atatgacaat	cctgaagaaa	ccacaggttc	3240
catagaacta	atatcctgtc	tctctctctc	tctctctctc	tctctttttt	tttctttttt	3300
cctttttcca	tggaatctgg	gtgggagagg	atactgcggg	caccagaatg	ctaaagtttc	3360
ctaacttttt	gaagtttctg	tagttcatcc	ttaatcctga	cacccatgta	aatgtccaaa	3420
atgttgatct	tccactgcaa	atttcaaaaag	ccttgtcaat	ggtcaagcgt	gcagcttggt	3480
cagcggttct	ttctgaggag	cggacaccgg	gttacattac	taatgagagt	tgggtagaac	3540
tctctgagat	gtgttcagat	agtgtaatgt	ctacattctc	tgatgtagtt	aagtattttac	3600
agatgttaaa	tggagtattt	ttatttttatg	tatatactat	acaacaatgt	tcttttttgt	3660
tacagctatg	cactgtaaat	gcagccttct	tttcaaaaact	gctaaatttt	tcttaatacaa	3720
gaatattcaa	atgtaattat	gaggtgaaac	aattatttga	actaacata	tttagaagct	3780
gaacttactg	cttatatata	tttgatttga	aaaacaaaaa	gacagtgtgt	gtgtctgttg	3840
agtgcacaaa	gagcaaaatg	atgctttccg	cacatccatc	ccttaggtga	gcttcaatct	3900
aagcatcttg	tcaagaaata	tcttagtccc	ctaaagggtat	taaccacttc	tgcgatattt	3960
ttccacattt	tcttgctgct	tgtttttctt	tgaagtttta	tacactggat	ttgttagggg	4020
aatgaaattt	tctcatctaa	aattttttcta	gaagatatca	tgatttttatg	taaaagtctct	4080
caatgggttaa	ccatttaagaa	atgttttttat	tttctctatc	aacagtagtt	ttgaaactag	4140
aagtcaaaaa	tcttttttaaa	atgctgtttt	gttttaattt	ttgtgatttt	aatttgatac	4200
aaaatgctga	ggtaataatt	atagtatgat	ttttacaata	attaatgtgt	gtctgaagac	4260
tatctttgaa	gccagtattt	ctttcccttg	gcagagtatg	acgatgggtat	ttatctgtat	4320
tttttacagt	tatgcacctc	gtataaatac	tgatatttca	ttccttttgt	tactaaagag	4380
acatatttat	cagttgcaga	tagcctattt	attataaatt	atgagatgat	gaaaataata	4440
aagccagtg	aaattttcta	cctaggatgc	atgacaattg	tcaggttgga	gtgtaagtgc	4500
ttcatttggg	aaattcagct	tttgacagaag	cagtgtttct	acttgcaact	gcattggcctc	4560
tgacgtgacc	atggtgttgt	tcttgatgac	attgctctctg	ctaaatttaa	taaaaacttc	4620
agaaaaacct	ccattttgat	catcaggatt	tcatctgagt	gtggagtccc	tggaaatgga	4680
ttcagtaaca	tttggagtgt	gtattcaagt	ttctaaattg	agattcgatt	actgtttggc	4740


```

tgacatgact tttctggaag acatgataca cctactactc aattgttctt ttcctttctc 4800
tcgccaaca cgatcttgta agatggattt caccocccagg ccaatgcagc taattttgat 4860
agctgcattc atttatcacc agcatattgt gttctgagtg aatccactgt ttgtcctgtc 4920
ggatgcttgc ttgatttttt ggcttcttat ttctaagtag atagaaagca ataaaaatac 4980
tatgaaatga aagaacttgt tcacagggtc tgcgttacaa cagtaacaca tctttaatcc 5040
gcctaattct tgttggtctg taggttaaatt gcagggtattt taactgtgtg aacgccaac 5100
taaagtttac agtctttctt tctgaatttt gagtatcttc tgttgtagaa taataataaa 5160
aagactatta agagcaataa attattttta agaaatcgag atttagtaaa tcctattatg 5220
tgttcaagga ccacatgtgt tctctatttt gcctttaaatt ttttgtgaac caatttttaa 5280
tacattctcc tttttgccct ggattgttga catgagtggg atacttggtt tcttttctta 5340
cttatcaaaa gacagcacta cagatatcat attgaggatt aatttatccc ccctaccccc 5400
agcctgacaa atattgtttac catgaagata gttttcctca atggacttca aattgcatct 5460
agaattagtg gagcttttgt atcttctgca gacactgtgg gtagcccatc aaaatgtaag 5520
ctgtgctcct ctcattttta tttttatttt tttgggagag aatatttcaa atgaacacgt 5580
gcaccccatc atcactggag gcaaatttca gcatagatct gtaggatttt tagaagaccg 5640
tgggcatatg ccttcacatg gtggtaagta ccacatctac aattttggta accgaactgg 5700
tgcttagta atgtggattt ttttcttttt taaaagagat gtagcagaat aattcttcca 5760
gtgcaacaaa atcaattttt tgctaaacga ctccgagaac aacagttggg ctgtcaacat 5820
tcaaagcagc agagagggaa ctttgacta ttgggggatg atgtttgggt cagttgataa 5880
aaggaaacct tttcatgcct ttagatgtga gcttccagta ggtaatgatt atgtgtcctt 5940
tcttgatggc tgtaatgaga acttcaatca ctgtagtcta agacctgac tatagatgac 6000
ctagaatagc catgtactat aatgtgatga ttctaaattt gtacctatgt gacagacatt 6060
ttcaataatg tgaactgctg atttgatgga gctactttaa gatttgtagg tgaaagtgt 6120
atactgttgg ttgaactatg ctgaagagg aaagtgagc attagttgag cccttgccgg 6180
gccttttttc cacctgcaa ttctacatgt attgttgtgg ttttattcat tgtatgaaa 6240
ttcctgtgat tttttttaaa tgtgcagtac acatcagcct cactgagcta ataaaggaa 6300
acgaatgttt caaatct 6317

```

<210> 232

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013261

<400> 232

ctgtagtcta agacctgac tatagatacc tagaatagcc atgtactata atgtgatgat 60

<210> 233

<211> 3237

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013277

<400> 233

```

gcgaagtga ggggtggcca ggtggggcca ggctgactga atgtatctcc tagctatgga 60
ctaaataata catgggggga aataaacaag tattcatgag ggtgaaaatg tgaccagca 120
ggaaaattac aactattttc aattgacgtt gaataggatg agtcatggaa tttaagtgt 180
ttactgaaga ttatactact ggtagataga agagctaaag aaagatggat actatgatgc 240
tgaatgtgcg gaatctgttt gagcagcttg tgcgcgggt ggagattctc agtgaaggaa 300
atgaagtcca atttatccag ttggcgaagg actttgagga tttccgtaaa aagtggcaga 360
ggactgacca tgagctgggg aaatacaagg atcttttgat gaaagcagag actgagcgaa 420
gtgctctgga tgtaagctg aagcatgcac gtaatcaggt ggatgtagag atcaaacgga 480
gacagagagc tgaggctgac tgcgaaaagc tggaacgaca gattcagctg attcgagaga 540
tgctcatgtg tgacacatct ggcagcattc aactaagcga ggagcaaaaa tcagctctgg 600
cttttctcaa cagaggccaa ccattccagca gcaatgctgg gaacaaaaga ctatcaacca 660
ttgatgaatc tggttccatt ttatcagata tcagctttga caagactgat gaatcactgg 720
attgggactc ttctttggtg aagactttca aactgaagaa gagagaaaag aggcgctcta 780

```

ctagccgaca	gtttgttgat	ggtccccctg	gacctgtaaa	gaaaactcgt	tccattggct	840
ctgcagtaga	ccaggggaat	gaatccatag	ttgcaaaaac	tacagtgact	gttcccaatg	900
atggcgggccc	catcgaagct	gtgtccacta	ttgagactgt	gccatattgg	accaggagcc	960
gaaggaaaac	aggtacttta	caaccttgga	acagtgactc	cacctgaac	agcaggcagc	1020
tggagccaag	aactgagaca	gacagtgtgg	gcacgccaca	gagtaatgga	gggatgcgcc	1080
tgcattgactt	tgtttctaag	acggttatta	aacctgaatc	ctgtgttcca	tgtggaaagc	1140
ggataaaatt	tggcaaatta	tctctgaagt	gtcgagactg	tcgtgtggtc	tctcatccag	1200

aatgtcggga	ccgctgtccc	cttccctgca	ttcctaccct	gataggaaca	cctgtcaaga	1260
------------	------------	------------	------------	------------	------------	------

ttggagaggg	aatgctggca	gactttgtgt	cccagacttc	tccaatgatc	ccctccattg	1320
ttgtgcattg	tgtaaatgag	attgagcaaa	gaggtctgac	tgagacaggc	ctgtatagga	1380
tctctggctg	tgaccgcaca	gtaaaagagc	tgaaagagaa	attcctcaga	gtgaaaactg	1440
tacccctcct	cagcaaagtg	gatgatatcc	atgctatctg	tagccttcta	aaagactttc	1500
ttcgaaacct	caaagaacct	cttctgacct	ttcgccttaa	cagagccttt	atggaagcag	1560
cagaaatcac	agatgaagac	aacagcatag	ctgccatgta	ccaagctgtt	ggtgaactgc	1620
cccaggccaa	cagggacaca	ttagctttcc	tcatgattca	cttgacagaga	gtggctcaga	1680
gtccacatac	taaaatggat	gttgccaatc	tggctaaagt	ctttggccct	acaatagtgg	1740
cccatgctgt	gcccaatcca	gacccagtga	caatgttaca	ggacatcaag	cgtcaaccca	1800
aggtgggtga	gcgcctgctt	tccttgccctc	tggagtattg	gagtcagttc	atgatgggtg	1860
agcaagagaa	cattgacccc	ctacatgtca	ttgaaaactc	aaatgccttt	tcaacaccac	1920
agacaccaga	tattaaagtg	agtttactgg	gacctgtgac	cactcctgaa	catcagcttc	1980
tcaagactcc	ttcatctagt	tccctgtcac	agagagtccg	ttccaccctc	accaagaaca	2040
ctcctagatt	tgggagcaaa	agcaagtctg	ccactaacct	aggacgacaa	ggcaactttt	2100
ttgcttctcc	aatgctcaag	tgaagtcaca	tctgcctgtt	acttcccagc	attgactgac	2160
tataagaaaag	gacacatctg	tactctgctc	tgcagcctcc	tgtactcatt	actactttta	2220
gcattctcca	ggcttttact	caagttaaat	tgtgcatgag	ggttttatta	aaactatata	2280
tatctcccct	tccttctcct	caagtcacat	aatatcagca	ctttgtgctg	gtcattgttg	2340
ggagctttta	gatgagacat	ctttccaggg	gtagaagggt	tagtatggaa	ttggttgtga	2400
ttcttttttg	ggaagggggg	tattgttcct	ttggcttaaa	gccaaatgct	gctcatagaa	2460
tgatctttct	ctagtttcat	ttagaactga	tttccgtgag	acaatgacag	aaaccctacc	2520
tatctgataa	gattagcttg	tctcaggggtg	ggaagtggga	gggcagggca	aagaaaggat	2580
tagaccagag	gatttaggat	gcctccttct	aagaaccaga	agttctcatt	ccccattatg	2640
aactgagcta	taatatggag	ctttcataaa	aatgggatgc	attgaggaca	gaactagtga	2700
tgggagtatg	cgtagctttg	atttggtatga	ttaggtcttt	aatagtgttg	agtggcacaa	2760
ccttgtaaat	gtgaaagtac	aactcgtatt	tatctctgat	gtgccgctgg	ctgaactttg	2820
ggttcatttg	gggtcaaagc	cagtttttct	tttaaaattg	aattcattct	gatgcttggc	2880
ccccataccc	ccaaccttgt	ccagtggagc	ccaacttcta	aaggtaata	tatcatcctt	2940
tggcatccca	actaacaata	aagagttagc	tataaggga	gattgtcaat	attttgtggt	3000
aagaaaagct	acagtcattt	tttctttgca	ctttggatgc	tgaaattttt	cccatggaac	3060
atagccacat	ctagatagat	gtgagctttt	tcttctgtta	aaattattct	taatgtctgt	3120
aaaaacgatt	ttcttctgta	gaatgtttga	cttcgtattg	acccttatct	gtaaaacacc	3180
tatttgggat	aatatattgga	aaaaaagtaa	atagcttttt	caaatgaaa	aaaaaaa	3237

<210> 234

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013277

<400> 234

ctcattcccc attatgaact gagctataat atggagcttt cataaaaaatg ggatgcattg 60

<210> 235

<211> 1122

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013409

<400> 235

```

gctcctcgcc ccgcgcctgc ccccaggatg gtccgcgcga ggcaccagcc ggggtgggctt 60
tgcctcctgc tgcctgctgc ctgccagttc atggaggacc gcagtgccca ggctgggaac 120
tgcctggctcc gtcaagcgaa gaacggccgc tgcagggtcc tgtacaagac cgaactgagc 180
aaggaggagt gctgcagcac cggccggctg agcacctcgt ggaccgagga ggacgtgaat 240
gacaacacac tcttcaagtg gatgattttc aacgggggcg cccccaactg catccctgt 300
aaagaaacgt gtgagaacgt ggactgtgga cctgggaaaa aatgccgaat gaacaagaag 360
aacaaccccc gctgcgtctg cgccccggat tgttccaaca tcacctgga ggggccagtc 420
tgcgggctgg atgggaaaac ctaccgcaat gaatgtgcac tcctaaaggc aagatgtaaa 480
gagcagccag aactggaagt ccagtaccaa ggcagatgta aaaagacttg tcgggatgtt 540
ttctgtccag gcagctccac atgtgtggtg gaccagacca ataatgccta ctgtgtgacc 600
tgtaatcgga tttgccaga gctgcttcc tctgagcaat atctctgtgg gaatgatgga 660
gtcacctact ccagtgcctg ccacctgaga aaggctacct gcctgctggg cagatctatt 720
ggattagcct atgagggaaa gtgtatcaaa gcaaagtcct gtgaagatat ccagtgcact 780
gggtgggaaaa aatgtttatg ggatttcaag gttgggagag gccggtgttc cctctgtgat 840
gagctgtgcc ctgacagtaa gtcggatgag cctgtctgtg ccagtgacaa tgccacttat 900
gccagcgagt gtgccatgaa ggaagctgcc tgcctcctcag gtgtgctact ggaagtaaa 960
cactccggat cttgcaactc catttcggaa gacaccgagg aagaggagga agatgaagac 1020
caggactaca gctttcctat atcttctatt ctagagtggg aaactctcta taagtgttca 1080
gtgttcacat agcctttgtg caaaaaaaaa aaaaaaaaaa aa 1122

```

<210> 236

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_013409

<400> 236

```

gaagatgaag accaggacta cagctttcct atatcttcta ttctagagtg gtaaactctc 60

```

<210> 237

<211> 11389

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014246

<400> 237

```

atggcgccgc cgccgcgcgc cgtgctgccc gtgctgctgc tcttggccgc cgccgcgcgc 60
ctgccgcgga tggggctgcg agcggcgcgc tgggagccgc gcgtaccggg cgggaccgc 120
gccttcgccc tccggcccggt ctgtacctac gcgggtggcg ccgcttgac gccccggcg 180
ccgcgggagc tgcctggacgt gggccgcgat gggcggtgg caggacgtcg gcgcgtctcg 240
ggcgcggggc gcccgctgcc gctgcaagtc cgcttggtgg cccgcagtgc cccgacggcg 300
ctgagccgcc gcctgcgggc gcgcacgcac ctcccggtc gcggagcccg tgcccggctc 360
tgccgaaccg gtgcccggct ctgcggggcg ctctgcttcc ccgtcccgg cggctgcgcg 420
gccgcgcagc attcggcgct cgagctccg accaccttac ccgcctgccc ctgcccgcgc 480
cgcccaggc cccgctgtcc cggccgtccc atctgcctgc cgccggcgcg ctccgtccgc 540
ctgcgtctgc tgtgcgccct gcggcgcccg gctggcgccg tccgggtggg actggcgctg 600
gaggccgcca ccgcggggac gccctcccg tcccatccc catcgccgc cctgccgcgc 660
aacttgcccg aagcccgggc ggggcggcg cgacggccc ggcggggcac gagcgcgaga 720
gggagcctga agtttccgat gcccaactac cagggtggcg tgtttgagaa cgaaccggcg 780
ggcacccctca tctccagct gcacgcgcac tacaccatcg agggcgagga ggagcgctg 840
agctattaca tggaggggct gttcgacgag cgctccggg gctacttccg aatcgactct 900
gccacggcg ccgtgagcac ctggaccgcg agaccaagga gacgcacgtc 960
ctcagggtga aagccgtgga ctacagtac cgccgcgcgt cggccaccac ctacatcact 1020

```

gtcttgggtca	aagacaccaa	cgaccacagc	ccggttcttcg	agcagtcgga	gtaccgagag	1080
cgctgtcggg	agaacctgga	ggtgggctac	gaggtgctga	ccatccgcgc	cagcgacgcg	1140
gactcgccca	tcaacgccaa	cttgcgttac	cgctgttg	ggggcgctg	ggacgtcttc	1200
cagctcaacg	agagctctgg	cgtgggtgagc	acacgggcgg	tgctggaccg	ggaggaggcg	1260
gccgagtacc	agctcctggt	ggaggccaac	gaccaggggc	gcaatccggg	cccgcctcagt	1320
gccacggcca	ccgtgtacat	cgaggtggag	gacgagaacg	acaactacc	ccagttcagc	1380
gagcagaact	acgtggatcca	ggtgcccag	gacgtggggc	tcaacaaggc	tgtgctgca	1440
gtgcaggcca	cggaccggga	ccagggccag	aacggggcca	ttactacag	catcctcagc	1500
gggaacgtgg	ccggccagtt	ctacctgcac	tcgctgagcg	ggatcctgga	tgtgatcaac	1560
cccttgatt	tcgaggatgt	ccagaaatac	tcgctgagca	ttaaggcca	ggatgggggc	1620
cggcccccg	tcatacaattc	ttcaggggtg	gtgtctgtgc	aggtgctgga	tgtcaacgac	1680
aacgagccta	tctttgtgag	cagccccctc	caggccacgg	tgctggagaa	tgtgccccctg	1740
ggctaccccc	tggtgcacat	tcaggcgggtg	gacgcggact	ctggagagaa	cgccccgctg	1800
cactatcgcc	tggtgcacac	ggcctccacc	tttctggggg	gcggcagcgc	tgggcctaag	1860
aatcctgccc	ccacccctga	cttccccctc	cagatccaca	acagctccgg	ttggatcaca	1920
gtgtgtgccc	agctggaccg	cgaggaggtg	gagcaactaca	gcttcggggg	ggaggcgggtg	1980
gaccacggct	cgccccccat	gagctcctcc	accagcgtgt	ccatcacggg	gctggacgtg	2040
aatgacaacg	accgggtgtt	cacgcagccc	acctacgagc	ttcgtctgaa	tgaggatgcg	2100
gccgtgggga	gcagcgtgct	gaccctgcag	gcccgcgacc	gtgacgcca	cagtgtgatt	2160
acctaccagc	tcacaggcgg	caacacccgg	aaccgctttg	cactcagcag	ccagagaggg	2220
ggcgccctca	tcacctggc	gctacctctg	gactacaagc	aggagcagca	gtacgtgctg	2280
gcggtgacag	catccgacgg	cacacggtcg	cacatcgcc	atgtccta	caacgtcact	2340
gatgccaa	cccacaggcc	tgtctttcag	agctccatt	acacagtga	tgtcagtga	2400
gacaggcctg	tgggcacctc	cattgctacc	ctcagtcca	acgatgagga	cacaggagag	2460
aatgcccgca	tcacctacgt	gattcaggac	cccgtgccgc	agttccgcat	tgaccccgac	2520
agtggcacca	tgtacacat	gatggagctg	gactatgaga	accaggtcgc	ctacacgctg	2580
accatcatgg	cccaggacaa	cgcatcccg	cagaaatcag	acaccaccac	cctagagatc	2640
ctcatctctg	atgccaatga	caatgcacc	cagttcctgt	gggatttcta	ccagggttcc	2700
atctttgctg	atgctccacc	ctcgaccagc	atctccagg	tctctgccac	ggaccgggac	2760
tcaggctcca	atgggctct	gctgtacacc	ttccagggtg	gggacgacgg	cgatggggac	2820
ttctacatcg	agccacgtc	cggtgtgatt	cgcacccagc	gccggctgga	ccgggagaa	2880
gtggccgtgt	acaacctttg	ggctctggct	gtggatcggg	gcagtcacc	tccccctagc	2940
gcctcggtag	aaatccaggt	gaccatcttg	gacattaatg	acaatgccc	catgtttgag	3000
aaggacgaac	tggagctgtt	tgttgaggag	aacaacccag	tggggtcggt	ggtggcaag	3060
attcgtgcta	agcaccctga	tgaaggccct	aatgcccaga	tcattgtatca	gattgtggaa	3120
ggggacatgc	cgcatctctt	ccagctggac	gtgctcaacg	gggacctgcg	tgccatggtg	3180
gagctggact	ttgaggtecg	gcgggagtat	gtgctgggtg	tgacggccac	gtcggctccg	3240
ctggtgagcc	gagccacggg	gcacatcctt	ctcgtggacc	agaatgacaa	ccgcctgtg	3300
ctgcccgaact	tcagatcct	cttcaacaac	tatgtacca	acaagtcca	cagtttccc	3360
accggcgtga	tcggctgcat	cccggcccat	gaccccgacg	tgctcagacag	cctcaactac	3420
accttcgtgc	agggcaacga	gctgcgcctg	ttgctgctgg	accccgccac	gggcgaactg	3480
cagctcagcc	gcgacctgga	caacaacccg	ccgtggagg	cgctcatgga	ggtgtctgtg	3540
tctgatggca	tcacacgctg	cacggccttc	tgacacctgc	gtgtcaccat	catcacggac	3600
gacatgctga	ccaacagcat	cactgtccgc	ctggagaaca	tgtcccagga	gaagttcctg	3660
tccccgctgc	tggccctctt	cgtggagggg	gtggccgcgc	tgctgtccac	caccaaggac	3720
gacgtcttcg	tcttcaacgt	ccagaacgac	accgacgtca	gctccaacat	cctgaacgtg	3780
accttctcgg	cgctgctgcc	tggcggcgctc	cgcgccaggt	tcttcccgtc	ggaggacctg	3840
caggagcaga	tctacctgaa	tcggacgctg	ctgaccacca	tctccacgca	gcgcgtgctg	3900
cccttcgacg	acaacatctg	cctgcgcgag	ccctgcgaga	actacatgaa	gtgcgtgtcc	3960
gttctcgat	tcgacagctc	cgcgcccttc	ctcagctcca	ccacctgct	cttcgggccc	4020
atccacccca	tcaacggcct	gcgctgcgcg	tgcccgcgcg	gcttcaccgg	cgactactgc	4080
gagacggaga	tcgacctctg	ctactccgac	ccgtgcggcg	ccaacggccg	ctgccgcagc	4140
cgcgagggcg	gctacacctg	cgagtgtctc	gaggacttca	ctggagagca	ctgtgagggtg	4200
gatgcccgct	caggccgctg	tgccaacggg	gtgtgcaaga	acgggggcac	ctgcgtgaac	4260
ctgctcatcg	gcggttcca	ctgcgtgtgt	cctcctggcg	agtatgagag	gccctactgt	4320
gaggtgacca	ccaggagctt	cccgccccag	tccttcgtca	ccttcggggg	cctgagacag	4380
cgtctccact	tcacatctc	cctcacgttt	gcactcagg	aaaggaaagg	cttgcttctc	4440
tacaacggcc	gcttcaatga	gaagcacgac	ttcatcgccc	tggagatcgt	ggacgagcag	4500
gtgcagctca	ccttctctgc	aggcgagaca	acaacgacgc	tggcaccgaa	ggttcccag	4560
ggtgtgagtg	acgggcgggtg	gcactctgtg	caggtgcagt	actacaacaa	gcccataatt	4620
ggccacctgg	gcctgcccc	tgggcccgtcc	ggggaaaaga	tggccgtggt	gacagtggat	4680

gattgtgaca	caaccatggc	tgtgcgcttt	ggaaaggaca	tcgggaacta	cagctgcgct	4740
gcccagggca	ctcagaccgg	ctccaagaag	tccctggatc	tgaccggccc	tctactcctg	4800
gggggtgtcc	ccaacctgcc	agaagacttc	ccagtgcaca	accggcagtt	cgtgggctgc	4860
atgcggaacc	tgtcagtcga	cggcaaaaat	gtggacatgg	ccggattcat	cgccaacaat	4920
ggcaccgggg	aaggctgcgc	tgtctcgagg	aacttctgcg	atgggaggcg	gtgtcagaat	4980
ggaggcacct	gtgtcaacag	gtggaatatg	tatctgtgtg	agtgtccact	ccgattcggc	5040
gggaagaact	gtgagcaagc	catgcctcac	ccccagctct	tcagcgggtga	gagcgtcgtg	5100
tcctggagtg	acctgaacat	catcatctct	gtgccctggg	acctgggggt	catgttcogg	5160
acccggaagg	aggacagcgt	tctgatggag	gccaccagtg	gtgggcccac	cagctttcgc	5220
ctccagatcc	tgaacaacta	cctccagttt	gagggtgtcc	acggcccctc	cgatgtggag	5280
tcctgatgac	tgtccgggtt	gcgggtgacc	gacggggagt	ggcaccacct	gctgatcgag	5340
ctgaagaatg	ttaaggagga	cagtggagatg	aagcacctgg	tcaccatgac	cttggactat	5400
gggatggacc	agaacaaggc	agatatcggg	ggcatgcttc	ccgggtgac	ggtaaggagc	5460
gtggtggtcg	gaggcgcttc	tgaagacaag	gtctccgtgc	gccgtggatt	ccgaggctgc	5520
atgcacgaag	tgaggatggg	ggggacgccc	accaacgtcg	ccaccctgaa	catgaacaac	5580
gcactcaagg	tcagggtgaa	ggacggctgt	gatgtggacg	acccctgtac	ctcgagcccc	5640
tgtcccccca	atagccgctg	ccacgacgcc	tgggaggact	acagctgcgt	ctgtgacaaa	5700
gggtaccttg	gaataaactg	tgtggatgcc	tgtcacctga	acccctgcga	gaacatgggg	5760
gcctgcgtgc	gctcccccg	ctccccgcag	ggctacgtgt	gcgagtgtgg	gcccagtcac	5820
tacgggcccgt	actgtgagaa	caaaactcgac	cttccgtgcc	ccagaggctg	gtgggggaac	5880
cccgctctgtg	gaccctgcc	ctgtgccgtc	agcaaaggct	ttgatcccca	ctgtaataag	5940
accaacggcc	agtgcacaatg	caaggagaat	tactacaagc	tcctagcccc	ggacacctgt	6000
ctgcccctgg	actgcttccc	ccatggctcc	cacagccgca	cttgcgacat	ggccaccggg	6060
cagtgtgcct	gcaagcccgg	cgtcatcggc	cgccagtgc	accgctgcga	caacccgttt	6120
gccgaggtca	ccacgctcgg	ctgtgaagtg	atctacaatg	gctgtcccaa	agcatttgag	6180
gccggcatct	ggtggccaca	gaccaagttc	ggcgagccgg	ctgcggtgcc	atgccctaag	6240
ggatccggtt	gaaatgcggg	ccgacactgc	agcggggaga	agggctggct	gccccagag	6300
ctotttaact	gtaccacccat	ctccttcgtg	gacctcaggg	ccatgaatga	gaagctgagc	6360
cgcaatgaga	cgcaggtgga	cggcgccagg	gcctgcagc	tggtgagggc	gctgcgcagt	6420
gctacacagc	acacgggcac	gctctttggc	aatgacgtgc	gcacggccta	ccagctctg	6480
ggccacgtcc	ttcagcacga	gagctggcag	cagggtctcg	acctggcagc	cacgcaggac	6540
gccgactttc	acgaggacgt	catccactcg	ggcagcgccc	tectggcccc	agccaccagg	6600
gcggcggtgg	agcagatcca	gcggagcgag	ggcgccacgg	cacagctgct	ccggcgccctc	6660
gagggctaact	tcagcaacgt	ggcagcgaac	gtgcggcgga	cgtacctgcg	gcccctcgtc	6720
atcgtcaccg	ccaacatgat	tcttgctgtc	gacatctttg	acaagttaa	ctttacggga	6780
gccagggtcc	cgcgattcga	caccatccat	gaagagtcc	ccaggagct	ggagtcctcc	6840
gtctccttcc	gtcccgactt	cttcagacca	cctgaagaaa	aagaaggccc	cctgctgagg	6900
ccggctggcc	ggaggaccac	cccgacagac	acgcgcccgg	ggcctggcac	cgagagggag	6960
gccccgatca	gcaggcggag	gcgacacct	gatgacgctg	gccagttcgc	cgtcgctctg	7020
gtcatcattt	accgaccct	ggggcagctc	ctgcccagac	gctacgacct	cgaccgtcgc	7080
agcctccggt	tgctccaccg	gcccacatt	aataccccga	tggtgagcac	gctggtgtac	7140
agcgaggggg	ctccgctccc	gagaccctcg	gagaggcccg	tcttggtgga	gttcgccctg	7200
ctggagggtg	aggagcgaac	caagcctgtc	tgcgtgttct	ggaaccactc	cctggccggt	7260
ggtgggacgg	gaggtgtgtc	tgcccggggc	tgcgagctcc	tgtccaggaa	ccggacacat	7320
gtcgctgcc	agtgcagcca	cacagccagc	tttgcggtgc	tcatggatat	ctccaggcgt	7380
gagaacgggg	aggtcctgcc	tctgaagatt	gtcacctatg	ccgctgtgtc	cttgtcactg	7440
gcagccctgc	tggtggcctt	cgtcctcctg	agcctggtcc	gcattgctgc	ctccaacctg	7500
cacagcatto	acaagcacct	cggcgtggcg	ctcttcctct	ctcagctggt	gttcgtgatt	7560
gggatcaacc	agacggaaaa	ccggtttctg	tgcacagtgg	ttgccatcct	cctccactac	7620
atctacatga	gcacctttgc	ctggaccctc	gtggagagcc	tgcattgtcta	ccgcatgctg	7680
accgaggtgc	gcaacatcga	cacggggccc	atgcggttct	actacgtcgt	gggctggggc	7740
atcccgccca	ttgtcacagg	actggcggtc	ggcctggacc	cccagggtta	cggggaacccc	7800
gacttctgct	ggctgtcgct	tcaagacacc	ctgatttgga	gctttgcggg	gcccacggga	7860
gctgttataa	tcatcaaac	agtcacttct	gtcctatctg	caaaggtttc	ctgcccaga	7920
aagcaccatt	attatgggaa	aaaagggatc	gtctccctgc	tgaggaccgc	attcctcctg	7980
ctgtgctca	tcagcgccac	ctggctgctg	gggctgctgg	ctgtgaaccg	cgatgactg	8040
agctttcaat	acctttcgc	catcttcagc	ggcttacagg	gccccttcgt	cctccttttc	8100
cactgcctgc	tcaaccagga	ggtccggaag	cacctgaagg	gcgtgctcgg	cgggaggaag	8160
ctgcaacctg	aggactccgc	caccaccagg	gccaccctgc	tgacgcgctc	cctcaactgc	8220
aacaccacct	tcgggtgacg	gcctgacatg	ctgcgcacag	acttgggcga	gtccaccgcc	8280
tcgctggaca	gcatcgtcag	ggatgaagg	atccagaagc	tcggcgtgtc	ctctgggctg	8340

gtgaggggca	gccacggaga	gccagacgcg	tccctcatgc	ccaggagctg	caaggatccc	8400
cctggccacg	attccgactc	agatagcgag	ctgtccctgg	atgagcagag	cagctcttac	8460
gcctctcac	actcgtcaga	cagcgaggac	gatggggtgg	gagctgagga	aaaatgggac	8520
ccggccaggg	gcgccgtcca	cagcaccccc	aaaggggacg	ctgtggccaa	ccacgttccg	8580
gccggctggc	ccgaccagag	cctgggtgag	agtgcacagt	aggaccccag	cggcaagccc	8640
cgctgaagg	tggagaccaa	ggtcagcgctg	gagctgcacc	gcgaggagca	gggcagtcac	8700
cgtggagagt	accccccgga	ccaggagagc	gggggcgag	ccaggcttgc	tagcagccag	8760
ccccagagc	agaggaaagg	catcttgaaa	aataaagtca	cctacccgcc	gccgctgacg	8820
ctgacggagc	agacgctgaa	gggccggctc	cgggagaagc	tggccgactg	tgagcagagc	8880
cccacatcct	cgcgcacgtc	ttccctgggc	tctggcggcc	ccgactgcgc	catcacagtc	8940
aagagccctg	ggagggagcc	ggggcggtgac	caactcaacg	gggtggccat	gaatgtgcgc	9000
actgggagcg	cccaggccga	tggctccgac	tctgagaaac	cgtgaggcaa	gcccgtcacc	9060
ccacacaggc	tgggcatca	ccctcagacc	tggagccca	aggggccact	gcccttgaag	9120
tggagtgggc	ccagatgtg	gcggtcccca	tgggtggcagc	ccccgactg	atcatccaga	9180
cacaaaggtc	ttggttctcc	caggagctca	gggcctgtca	gacctggtga	caagtgcaca	9240
aggccacagg	catgagggag	gcgtggacca	ctgggccagc	accgctgagt	cctaagactg	9300
cagtcaaagc	cagaactgag	aggggacccc	agactgggcc	cagaggctgg	ccagagttca	9360
ggaacgccgg	gcacagacca	aagaccgcgg	tcagcccccg	cccaggcggg	catctcatgg	9420
cagtgcggac	ccgtggctgg	cagcccgggc	agtcccttgc	aaaggcacc	cttgtcttaa	9480
aatcacttcg	ctatgtggga	aaggtggaga	tacttttata	tatttgtatg	ggactctgag	9540
gaggtgcaac	ctgtatatat	attgcattcg	tgtgactttt	gttatcccga	gagatccatg	9600
caatgatctc	ttgtgtctct	ctctgtcaag	attgcacagt	tgtacttgaa	tctggcatgt	9660
gttgacgaaa	ctggtgcccc	agcagatcaa	aggtgggaaa	tacgtcagca	gtggggctaa	9720
aaccaagcgg	ctagaagccc	tacagctgcc	ttcggccagg	aagtgaggat	ggtgtgggcc	9780
ctccccgcgg	gccccctggg	tcccagtggt	tcgtgtgtgt	tgcgtttgtc	ctctgctgcc	9840
atctgccccg	gctgtgtgaa	ttcaagacag	ggcagtgacg	cactaggcag	gtgtgaggag	9900
ccctgctgag	gtcactgttg	ggcacggttg	ccacacggct	gtcatttttc	acctggtcat	9960
tctgtgacca	ccaccccctc	ccctcacgcg	ctcccagggt	gcccgggagc	tgcaggtggg	10020
gatggctttg	tcctttgtct	ctgtcccccg	tgggacctgg	gaccttaaa	cgttgcaggt	10080
tcctgatttg	gacagaggtg	tggggccttc	caggccgtta	catacctcct	gccaattctc	10140
taactctctg	agactgcgag	gatctccagg	cagggttctc	ccctctggag	tctgaccaat	10200
tacttcattt	tgttcaaat	ggccaattgt	gcagagggac	aaagccacag	ccacactctt	10260
caacggttac	caaactgttt	ttggaaatc	acaccaaggt	cgggcccact	gcaggcagct	10320
ggcacagcgt	ggcccagagg	gctgtggaac	gggtcccggg	actgtcagac	atgtttgatt	10380
ttagcgtttc	ctttgttctt	caaatcagggt	gcccataata	gtgatcagca	cagctgcttc	10440
caaataggag	aaaccataaa	ataggatgaa	aatcaagtaa	aatgcaaaga	tgtccacact	10500
gttttaaaact	tgaccctgat	gaaaatgtga	gcactgttag	cagatgccta	tgggagagga	10560
aaagcgtatc	tgaaaatggt	ccaggacagg	aggatgaaat	gagatcccag	agtcctcaca	10620
cctgaatgaa	ttatacatgt	gccttaccag	gtgagtggtc	tttcgaagat	aaaaaactct	10680
agtcccttta	aacgtttgcc	cctggcggtt	cctaagtagc	aaaaggtttt	taagtcttcg	10740
aacagtctcc	tttcatgact	ttaacaggat	tctgccccct	gaggtgtaat	ttttttgttc	10800
tatttttttc	cacgtactcc	acagccaaca	tcacgaggtg	taatttttaa	tttgatcaga	10860
actgttacca	aaaaacaact	gtcagtttta	ttgagatggg	aaaaatgtaa	acctattttt	10920
attacttaag	actttatggg	agagattaga	cactggaggt	ttttaacaga	acgtgtattt	10980
attaatgttc	aaaacactgg	aattacaaat	gagaagagtc	tacaataaat	taagattttt	11040
gaatttgtag	ttctgcgggt	ctggtttttc	tcacaaaaca	cccccgcccc	tccccatgcc	11100
cagggtggcc	gtggaaggga	cggtttacgg	acgtgcagct	gagctgtccg	tgtcccatgc	11160
tcctcagcc	agtggaaact	gccggaactt	tttgtocatt	ccctagtagg	cctgccacag	11220
cctagatggg	cagtttttgt	ctttcaccaa	atttgaggac	tttttttttt	tgccattatt	11280
tcttcagttt	tcttttcttg	cactgatctt	tctcctctcc	ttctgtgact	ccagtgactc	11340
agacgttaga	cctcttgatg	ttttcccact	ggtccctgag	gctctgttc	11389	

<210> 238

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014246

<400> 238
gggagagatt agacactgga gggtttttaac agaacgtgta tttattaatg ttcaaaacac 60

<210> 239
<211> 4372
<212> DNA
<213> Homo sapiens

<300>
<308> NM_014314

<400> 239
tagttattaa agttcctatg cagctccgcc tcgctccgg cctcatttcc tcggaaaatc 60
cctgctttcc ccgctcgcca cgccctcctc ctaccgggct ttaaagctag tgaggcacag 120
cctgcgggga ccgtagctag ctgcaagcag aggccggcat gaccaccgag cagcgacgca 180
gcctgcaagc cttccaggat tatatccgga agacctgga ccctacctac atcctgagct 240
acatggcccc ctggttttagg gaggaagagg tgcagtatat tcaggctgag aaaaacaaca 300
agggcccaat ggaggctgcc acactttttc tcaagttcct gttggagctc caggaggaag 360
gctggttccg tggctttttg gatgccctag accatgcagg ttattctgga ctttatgaag 420
ccattgaaag ttgggatttc aaaaaattg aaaagtggga ggagtataga ttacttttaa 480
aacgtttaca accagaattt aaaaccagaa ttatcccaac cgatatcatt tctgatctgt 540
ctgaatgttt aattaatcag gaatgtgaag aaattctaca gatttgctct actaagggga 600
tgatggcagg tgcagagaaa ttggtggaat gccttctcag atcagacaag gaaaactggc 660
ccaaaacttt gaaacttgct ttggagaaag aaaggaacaa gttcagtga ctgtggattg 720
tagagaaagg tataaaagat gttgaaacag aagatcttga ggataagatg gaaacttctg 780
acatacagat tttctaccaa gaagatccag aatgccagaa tcttagtgag aattcatgtc 840
caccttcaga agtgtctgat acaaacttgt acagcccatt taaaccaaga aattaccaat 900
tagagcttgc tttgcctgct atgaaaggaa aaaacacaa aatatgtgct cctacagggt 960
gtggaaaaac ctttgtttca ctgcttatat gtgaacatca tcttaaaaaa tcccacaag 1020
gacaaaagg gaaagtgttc ttttttgca atcagatccc agtgatgaa cagcagaaat 1080
ctgtattctc aaaatacttt gaaagacatg ggtatagagt tacaggcatt tctggagcaa 1140
cagctgagaa tgtcccagtg gaacagattg ttgagaacaa tgacatcatc attttaactc 1200
cacagattct tgtgaacaac cttaaaaagg gaacgattcc atcactatcc atctttactt 1260
tgatgatatt tgatgaatgc cacaacacta gtaaacacaa cccgtacaat atgatcatgt 1320
ttaattatct agatcagaaa cttggaggat cttcaggccc actgccccag gtcattgggc 1380
tgactgcctc ggttggtgtt ggggatgcca aaaaacacaga tgaagccttg gattatatct 1440
gcaagctgtg tgcttctctt gatgcgtcag tgatagcaac agtcaaacac aatctggagg 1500
aactggagca agttgtttat aagccccaga agtttttcag gaaagtggaa tcacggatta 1560
gcgacaaatt taaatacatc atagctcagc tgatgaggga cacagagagt ctggcaaaga 1620
gaatctgcaa agacctcgaa aacttatctc aaattcaaaa tagggaattt ggaacacaga 1680
aatatgaaca atggattgtt acagttcaga aagcatgcat ggtgttccag atgccagaca 1740
aagatgaaga gagcaggatt tgtaaagccc tgtttttata cacttcacat ttgcggaat 1800
ataatgatgc cctcattatc agtgagcatg cacgaatgaa agatgctctg gattacttga 1860
aagacttctt cagcaatgtc cgagcagcag gattcgatga gattgagcaa gatcttactc 1920
agagatttga agaaaagctg caggaactag aaagtgttcc cagggatccc agcaatgaga 1980
atcctaaact tgaagacctc tgcttcactc tacaagaaga gtaccactta aaccagaga 2040
caataacaat tctctttgtg aaaaccagag cacttgtgga cgctttaaaa aattggattg 2100
aaggaaatcc taaactcagt tttctaaaac ctggcatatt gactggacgt ggcaaaacaa 2160
atcagaacac aggaatgacc ctcccggcac agaagtgtat attggatgca ttcaaagcca 2220
gtggagatca caatattctg attgccacct cagttgctga tgaaggcatt gacattgcac 2280
agtcaatct tgtcatcctt tatgagtatg tgggcaatgt catcaaaatg atccaaacca 2340
gaggcagagg aagagcaaga ggtagcaagt gcttcttct gactagtaat gctggtgtaa 2400
ttgaaaaaga acaataaac atgtacaaag aaaaaatgat gaatgactct attttacgcc 2460
ttcagacatg ggacgaagca gtatttaggg aaaagattct gcatatacag actcatgaaa 2520
aattcatcag agatagtcaa gaaaaaccaa aacctgtacc tgataaggaa aataaaaaac 2580
tgctctgcag aaagtgc aaa gcttggcat gttacacagc tgacgtaaga gtgatagagg 2640
aatgccatca cactgtgctt ggagatgctt ttaaggaaatg ctttgtgagt agaccacatc 2700
ccaagccaaa gcagtgtttca agttttgaaa aaagagcaaa gatattctgt gcccgacaga 2760
actgcagcca tgactgggga atccatgtga agtacaagac atttgagatt ccagttataa 2820
aaattgaaag ttttgggtg gaggatattg caactggagt tcagacactg tactcgaagt 2880
ggaaggactt tcattttgag aagataccat ttgatccagc agaaatgtcc aaatgatatc 2940

```

aggtcctcaa tcttcagcta caggggaatga gtaactttga gtggagaaga aacaaacata 3000
gtgggtataa tcatggatcg cttgtacccc tgtgaaaata tattttttta aaatatcttt 3060
agcagtttgt actatattat atatgcaaag cacaaatgag tgaatcacag cactgagtat 3120
ttttagggcc aacagagctc atagtacttg ggaaaaatta aaaagcctca tttctagcct 3180
tctttttaga gtcaactgcc aacaaacaca cagtaatcac tctgtacaca ctgggataga 3240
tgaatgaatg gaatgttggg aatttttatc tccctttgtc tccttaacct actgtaaact 3300
ggcttttgcc cttaacaatc tactgaaatt gttcttttga aggttaccag tgactctggt 3360
tgccaaatcc actgggcact tottaacctt ctatttgacc tctgcgcatt tggccctggt 3420
gagcactctt ctggaagctc tccctgggct tctctctctt ctagtcttat tctagtcttt 3480
ttttattgag tctctctctt tgctgatccc ttccaagggt tcaatatata tacatgtata 3540
tactgtacat atgtatatgt aactaatata catacataca ggtatgtata tgtaatggtt 3600
atatgtactc atgttctctg tgtagcaacg tgtggtatgg ctacacagag aacatgagaa 3660
cataaagcca tttttatgct tactactaaa agctgtccac tgtagagttg ctgtatgtag 3720
caatgtgtat ccactctaca gtggtcagct tttagtagag agcataaaaa tgataaaaata 3780
cttcttgaaa acttagttta ctatacatct tgccctatta atatgttctc ttaacgtgtg 3840
ccattgttct ctttgacctt ttctctataa tgatgttgat gttcaacacc tggactgaat 3900
gtctgttctc agatcccttg gatgttacag atgaggcagt ctgactgtcc tttctacttg 3960
aaagattaga atatgtatcc aaatggcatt cacgtgtcac ttagcaagggt ttgctgatgc 4020
ttcaaagagc ttagtttgcg gtttcttgga cgtggaaaca agtatctgag ttccctggag 4080
atcaacggga tgaggtgtta cagctgcctc cctcttcatg caatctggtg agcagtggtg 4140
caggcgggga gccagagaaa cttgccagtt atataacttc tctttggcct ttcttcatct 4200
gtaaaacaag gataatactg aactgtaagg gttagtggag agtttttaat taaaagaatg 4260
tgtgaaaagt acatgacaca gtagttgctt gataatagtt actagtagta gtattcttac 4320
taagacccaa tacaaatgga ttatttaaac caaaaaaaaa aaaaaaaaaa aa 4372

```

<210> 240

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014314

<400> 240

```

agttcagaca ctgtactcga agtgggaagga ctttcatttt gagaagatac catttgatcc 60

```

<210> 241

<211> 1647

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014321

<400> 241

```

gcgcgcgggt ttcggtgacc cgcggcggtc acgggaattg ttcgcttttag tgccggcgcc 60
atggggctcg agctgatcgg gcgcctagcc ccgcgcctgg gcctcgccga gcccgacatg 120
ctgaggaaaag cagaggagta cttgcgcctg tcccggttga agtggtgctg cctctccgca 180
cgcaccacgg agaccagcag tgcagtcag tgccctggacc ttgcagcttc ctggatgaag 240
tgcccccttg acagggtcta tttaattaaa ctttctgggt tgaacaagga gacatatcag 300
agctgtctta aatcttttga gtgtttactg ggctgaatt caaatatttg aataagagac 360
ctagctgtac agtttagctg tatagaagca gtgaacatgg cttcaaagat actaaaaagc 420
tatgagtcca gtcttcccca gacacagcaa gtggatcttg acttatccag gccacttttc 480
acttctgctg cactgctttc agcatgcaag attctaaagc tgaaagtggg taaaaacaaa 540
atggtagcca catccggtgt aaaaaaagct atatttgatc gactgtgtaa acaactagag 600
aagattggac agcaggctga cagagaacct ggagatgtag ctactccacc acggaagaga 660
aagaagatag tggttgaagc ccagcaaaag gaaatggaga aggtagagga gatgccacat 720
aaaccacaga aagatgaaga tctgacacag gattatgaag aatggaaaag aaaaattttg 780
gaaaatgctg ccagtgtca aaaggctaca gcagagtgat ttcagcttcc aaactggtat 840
acattccaaa ctgatagtag attgccatct ccagggaagc ttgacggcct tgggattttg 900

```



```

tttaaaacttt tataataagg atcctaagac tgttgccctt aaatagcaaa gcagccctacc 960
tggaggctaa gtctgggcag tgggctggcc cctggtgtga gcattagacc agccacagt 1020
cctgattggt atagccttat gtgctttcct acaaaatgga attggaggcc gggcgagtg 1080
gctcacgcct gtaatcccag cactttggga ggccaagggt ggtggatcac ctgaggtcag 1140
gagctcgaga ccagcctggc caacatggtg aaaccccatc tctactaaaa atacaaaaat 1200
tagccagggt tgatggtgca tgcctgtaat cccagctcct cagtaggctg agacaggagc 1260
atcacttgaa cgtgggaggc agaggttgca gtgagccgag attgcaccac cgactccag 1320
cctgggtgac agagcgagac ttatctcata aataaataga tagatactcc agcctgggtg 1380
acagagcgag acttatagat agatagatag atagatggat agatagatag atagatagat 1440
agatagataa acggaattgg agccattttg ctttaagtga atggcagtc cttgtcttat 1500
tcagaatata aaattcagtc tgaatggcat cttacagatt ttacttcaat ttttgtgtac 1560

```

```

ggtatttttt atttgactaa atcaatatat tgtacagcct aagttaataa atgttattta 1620
tatatgcaaa aaaaaaaaaa aaaaaaa 1647

```

```

<210> 242
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_014321

```

```

<400> 242
tgctttaagt gaatggcagt cccttgtctt attcagaata taaaattcag tctgaatggc 60

```

```

<210> 243
<211> 1455
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_014364

```

```

<400> 243
ggcgggtccgc acgcacctcg gtaacatcac agcagggtcca ggccaatgat aaccttataa 60
gagggccatgt cgaagcgaga catcgtccct accaatgtca ccgttgtcca gttgctgcga 120
cagccgtgcc cgggtgaccag agcacgcgcc ccacctgagc ctaaggctga agtagagccc 180
cagccacaac cagagcccac accagtcagg gaggaataa agccaccacc gccaccactg 240
cctcctcacc ccgctaactc tcctcctaag atggtgtctg tggcccgga gctgactgtg 300
ggcatcaatg gatttggacg catcggtgcg ctggtcctgc gcgcctgcat ggagaagggt 360
gttaagggtg tggtgtgaa tgatccattc attgaccggg aatacatggt gtacatgttt 420
aagtatgact ccacccacgg ccgatacaag ggaagtgtgg aattcaggaa tggacaactg 480
gtcgtggaca accatgagat ctctgtctac cagtgc aaag agcccaaaca gatcccctgg 540
agggctgtcg ggagccccta cgtggtggag tccacaggcg tgtacctc catacaggca 600
gcttcggacc acatctctgc aggtgctcaa cgtgtggtca tctccgcgcc ctcaccggat 660
gcaccaatgt tcgtcatggg tgtcaatgaa aatgactata accctggctc catgaacatt 720
gtgagcaacg cgtcctgcac caccaactgt ttggtcctcc tcgccaaagt catccacgag 780
cgatttggga tcgtggaagg gttgatgacc acagtccatt cctacacggc caccagaag 840
acagtggacg ggccatcaag gaaggcctgg cgagatgggc ggggtgccca ccagaacatc 900
atcccagcct ccactggggc tgcgaaagct gtgaccaaag tcatcccaga gctcaaagg 960
aagctgacag ggatggcggt ccgggtacca acccggatg tgtctgtcgt ggacctgacc 1020
tgccgcctcg ccagcctgc cccctactca gccatcaagg aggtgtgaaa agcagcagcc 1080
aaggggccc aaggctggcat ccttgcttac accgaggatg aggtcgtctc tacggacttc 1140
ctcgggtgata cccactcgtc catcttcgat gctaaggccg gcattgcgct caatgacaat 1200
ttcgtgaagc tcatttcatt gtacgacaac gaatatggct acagtcaccg ggtggctgac 1260
ctcctccgct acatgttcag ccgagacaag tgaaacggga aggtcctttc tttccttccc 1320
aggggcgggg gccggaacat gtgcctcccg ttccagcatc tggctgccc ggggaggaag 1380
gacaccggg gcgggcgccc cagccgagtg ggtccatggt gaaataaaaa acagtgtctg 1440
aaaaaaaaa aaaaa 1455

```

<210> 244
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_014364

<400> 244
 cgctcaatga caatttcgtg aagctcattt catgggtacga caacgaatat ggctacagtc 60

<210> 245
 <211> 935
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_014462

<400> 245
 gaagtgggta agggtaatat ggaggagctt ccggcaggcc ccggcggctg aaagccgggg 60
 cagaagtgc ggtctcggtc gggattccgg gcttgggtccc accgaggcgg cgactgcggt 120
 aggaggggaag aggttttggga cgcgctggcc tcccgcgct gtgcattgca gcattatttc 180
 agttcaaaat gaactatatg cctggcaccg ccagcctcat cgaggacatt gacaaaaagc 240
 acttgggttct gcttcgagat ggaaggacac ttataggctt tttaagaagc attgatcaat 300
 ttgcaaactt agtgctacat cagactgtgg agcgtattca tgtgggcaaa aaatacgggtg 360
 atattcctcg agggattttt gtggtcagag gagaaaatgt ggtcctacta ggagaaatag 420
 acttggaaaa ggagagtgac acaccctcc agcaagtatc cattgaagaa attctagaag 480
 aacaaagggt ggaacagcag accaagctgg aagcagagaa gttgaaagtg caggccctga 540
 aggaccgagg tctttccatt cctcgagcag atactcttga tgagtactaa tcttttgccc 600
 agaggctgtt ggctcttgaa gagtaggggc tgtcactgag tgaaagtgac atcctggcca 660
 cctcacgcat ttgatcacag actgtagagt ttgaaaagt cacttttatt tttaattatt 720
 ttacatatgc aacatgaaga aatcgtgtag gtgggtttt tttttaataa caaaatcact 780
 gtttaaagaa acagtggcat agactccttc acacatcact gtggcaccag caactacttc 840
 tttatattgt tcttcatatc ccaaattaga gtttacaggg acagtcttca tttacttgta 900
 aataaaatat gaatctcaaa aaaaaaaaaa aaaaa 935

<210> 246
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_014462

<400> 246
 ttaataacaa aatcactgtt taaagaaaca gtggcataga ctcttcaca catcactgtg 60

<210> 247
 <211> 890
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_014501

<400> 247
 ggcggaccga agaacgcagg aagggggccg gggggaccgg ccccgggccg gccgcagcca 60
 tgaactccaa cgtggagaac ctacccccgc acatcatccg cctgggtgtac aaggaggtga 120
 cgacactgac cgcagaccga cccgatggca tcaaggtctt tcccaacgag gaggacctca 180
 ccgacctcca ggtcaccatc gagggccctg aggggacccc atatgctgga ggtctgttcc 240

```

gcatgaaact cctgctgggg aaggacttcc ctgcctcccc acccaagggc tacttcctga 300
ccaagatctt ccacccgaac gtgggcgcca atggcgagat ctgcgtcaac gtgctcaaga 360
gggactggac ggctgagctg ggcattccgac acgtactgct gaccatcaag tgcttctga 420
tccaccctaa ccccgagtct gcaactcaacg aggaggcggg ccgcctgctc ttggagaact 480
acgaggagta tgcggctcgg gcccgctctgc tcacagagat ccacggggggc gccggcggggc 540
ccagcggcag ggccgaagcc ggtcggggccc tggccagtgg cactgaagct tcctccaccg 600
accctggggc cccagggggc ccgggagggg ctgaggggtcc catggccaag aagcatgctg 660
gcgagcgcga taagaagctg gcggccaaga aaaagacgga caagaagcgg gcgctgcggg 720
cgctgcggcg gctgtagtgg gctctcttcc tccttccacc gtgaccccaa cctctcctgt 780
ccctccctc caactctgtc tctaagttat ttaaattatg gctggggctg gggagggtac 840
agggggcact gggacctgga tttgtttttc taaataaagt tggaaaagca 890

```

<210> 248

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014501

<400> 248

```

acacgtactg ctgaccatca agtgcctgct gatccaccct aaccccgagt ctgcactcaa 60

```

<210> 249

<211> 1182

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016095

<400> 249

```

gcggcgggcg gcgtctctc cggggacgct gaggggcccg aggagaccgt gaggtcttgg 60
cctgcagctc gcgcgcgccat ggacgctgcc gaggtcgaat tcctcgccga gaaggagctg 120
gttaccatta tccccaactt cagtctggac aagatctacc tcatcggggg ggacctgggg 180
ccttttaacc ctggtttacc cgtggaagtg cccctgtggc tggcgattaa cctgaaacaa 240
agacagaaat gtcgcctgct ccctccagag tggatggatg tagaaaagt ggagaagatg 300
agggatcatg aacgaaagga agaaaacttt accccaatgc ccagccctta ctacatggaa 360
cttacgaagc tcctgtttaa tcatgcttca gacaacatcc cgaaggcaga cgaaatccgg 420
acctgggtca aggatattg ggacactcgt atagccaaac tccgagtgtc tgctgacagc 480
tttgtgagac agcaggaggc acatgccaa ctggataact tgaccttgat ggagatcaac 540
accagcggga ctttctcac acaagcgctc aaccacatgt acaaaactcc caccgaacctc 600
cagcctctgg agagtactca gtctcaggac ttctagagaa aggcctgggtg caggcgggctt 660
gctgggggat gtgagcgctc aggatgtgat gaggtactcg tggttctgga gctctagaaa 720
cacttctgat gcatgaaaaa tgtgtgatgg tgcaaggaaat ggattcagga tgttgttgga 780
gaaacaagtt tgtgattagt ccttaaaact tagctccctg ggacattctt caattccaca 840
tctgtttcta gaaaccagcc ctttttcccc ccacttttga gaaataaaaa agccttaggt 900
aaataagtca ttctccctag cagagccact tgggtctcct gcatggaagc cgtcacactt 960
gggcagggtg tcagtactg gtaggtgtag atacagcagg agtggccatg tgggccacgg 1020
ctttttaccc cttcttgatc ctgatttctt gggctgaatt tagactctct cacagagggtg 1080

```

```

gctcacagag aaggatggca gatgggtgcag ccaacaatgc tgaccgggtg ttatcctcta 1140
agccctgatc cacaataaaa atggacccaa ctcaaaaaaa aa 1182

```

<210> 250

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016095

<400> 250
atggattcag gatgttgttg gagaaacaag tttgtgatta gtccttaaaa cttagctccc 60

<210> 251
<211> 704
<212> DNA
<213> Homo sapiens

<300>
<308> NM_016185

<400> 251
tgcagcgggtg gtcggctgtt ggggtgtggag tttcccagcg cccctcgggt ccgacccttt 60
gagcgttctg ctccggcgcc agcctacctc gtcctcggc gccatgacca caaccaccac 120
cttcaaggga gtcgaccca acagcaggaa tagctcccga gttttgcggc ctccagggtg 180
tgatccaat ttttcattag gttttgatga accaacagaa caacctgtga ggaagaacaa 240
aatggcctct aatatctttg ggacacctga agaaaatcaa gcttcttggg ccaagtcagc 300
aggtgccaag tctagtgttg gcagggaaga cttggagtca tctggactgc agagaaggaa 360
ctcctctgaa gcaagctccg gagacttctt agatctgaag ggagaagggtg atattcatga 420
aaatgtggac acagacttgc caggcagcct ggggcagagt gaagagaagc ccgtgcctgc 480
tgcgctgtg cccagcccgg tggccccggc cccagtgcc tccagaagaa atccccctgg 540
cggcaagtc agcctcgtct tgggttagct ctgactgtcc tgaacgctgt cgttctgtct 600
gtttcctcca tgcttgagaa ctgcacaact tgagcctgac tgtacatctt cttggatttg 660
tttcattaaa aagaagcact ttatgtaaaa aaaaaaaaaa aaaa 704

<210> 252
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> NM_016185

<400> 252
tgaaccaaca gaacaacctg tgaggaagaa caaaatggcc tctaatactt ttgggacacc 60

<210> 253
<211> 2268
<212> DNA
<213> Homo sapiens

<220>
<221> Modified_base
<222> 1 ... 2268
<223> n = a,c,g, or t

<300>

<308> NM_016359

<400> 253
gggatttgaa ccncgctgac gaagtttgggt gatccatctt ccgagtatcg ccgggatttc 60
gaatcgcat gatcatcccc tctctagagg agctggactc cctcaagtac agtgacctgc 120
agaacttagc caagagtctg ggtctccggg ccaacctgag ggcaaccaag ttgttaaaag 180
ccttgaaagg ctacattaaa catgaggcaa gaaaaggaaa tgagaatcag gatgaaagtc 240
aaacttctgc atcctcttgt gatgagactg agatacagat cagcaaccag gaagaagctg 300
agagacagcc acttggccat gtcaccaaaa caaggagaag gtgcaagact gtccgtgtgg 360
accctgactc acagcagaat cattcagaga taaaaataag taatccact gaattccaga 420
atcatgaaaa gcaggaaagc caggatctca gagctactgc aaaagtccct tctccaccag 480
acgagcacca agaagctgag aatgctgttt cctcaggtaa cagagattca aaggtacctt 540

```

cagaaggaaa gaaatctctc tacacagatg agtcatccaa acctggaaaa aataaaagaa 600
ctgcaatcac tactccaaac tttaagaagc ttcatgaagc tcattttaag gaaatggagt 660
ccattgatca atatataggag agaaaaagaa acatttttgaa gaacacaatt ccatgaatga 720
actgaagcag cagcccatca ataagggagg ggtcaggact ccagtacctc caagaggaag 780
actctctgtg gcttctactc ccatcagcca acgacgctcg caaggccggt cttgtggccc 840
tgcaagtcag agtaccttgg gtctgaaggg gtcaactcaag cgctctgcta tctctgcagc 900
taaaacgggt gtcaggtttt cagctgtctac taaagataat gagcataagc gtactactgac 960
caagactcca gccagaaagt ctgcacatgt gaccgtgtct gggggcacc caaaaggcga 1020
ggctgtgctt gggacacaca aattaaagac catcacgggg aattctgctg ctgttattac 1080
cccattcaag ttgacaactg aggcaacgca gactccagtc tccaataaga aaccagtgtt 1140
tgatcttaaa gcaagtttgt ctctgtcccct caactatgaa ccacacaaag gaaagctaaa 1200
accatggggg caatctaaag aaaataatta tctaaatcaa catgtcaaca gaattaactt 1260
ctacaagaaa acttacaac aaccccatct ccagacaaag gaagagcaac ggaagaaacg 1320
cgagcaagaa cgaaaggaga agaaagcaaa ggttttggga atgcgaaggg gcctcatttt 1380
ggctgaagat taataatttt ttaatatctt gtaaattatc ctgtattctc aacttttttc 1440
cttttgtaaa tttttttttt tttgctgtca tcccacttt agtcacgaga tctttttctg 1500
ctaactgttc atagtctgtg tagtgtccat gggttcttca tgtgctatga tctctgaaaa 1560
gacgttatca ccttaaagct caaattcttt gggatgggtt ttacttaagt ccattaacaa 1620
ttcagggttc taacgagacc catcctaaaa ttctgtttct agatttttaa tgtcaagttc 1680
ccaagttccc cctgctgggt ctaatattaa cagaactgca gtcttctgct agccaatagc 1740
atttacctga tggcagctag ttatgcaagc ttcaggagaa tttgaacaat aacaagaata 1800
gggtaagctg gtagagaaag gccacctctt cactctctat agaatatagt aacctttatg 1860
aaacggggcc atatagtttg gttatgacat caatatttta cctaggtgaa attgtttagg 1920
cttatgtacc ttctgtcaaa tatcctcatg taattgccat ctgtcactca ctatattcac 1980
aaaaataaaa ctctacaact cattctaaca ttgcttactt aaaagctaca tagccctatc 2040
gaaatgcgag gattaatgct ttaatgcttt tagagacagg gtctcactgt gttgccagg 2100
ctgggtctcaa actccaccaa atgtacttct tattcatttt atggaaaaga ctaggctttg 2160
cttagtatca tgtccatgtt tccttcacct cagtggagct tctgagtttt atactgctca 2220
agatcgtcat aaataaaatt ttttctcatt gtcaaaaaaa aaaaaaaa 2268

```

<210> 254

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016359

<400> 254

```

acattgctta cttaaaagct acatagccct atcgaaatgc gaggattaat gctttaatgc 60

```

<210> 255

<211> 1590

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016816

<400> 255

```

gaggcagttc tgttgccact ctctctctctg tcaatgatgg atctcagaaa taccocagcc 60
aaatctctgg acaagttcat tgaagactat ctcttgccag acacgtgttt ccgcatgcaa 120
atcgaccatg ccattgacat catctgtggg ttctggaagg aaaggtgctt ccgaggtagc 180
tccctaccctg tgtgtgtgtc caaggtggta aaggggtggc cctcaggcaa gggcaccacc 240
ctcagaggcc gatctgacgc tgacctgggt gtcttctcctc gtctctcac cacttttcag 300
gatcagttaa atcgccgggg agagttcato caggaaatta ggagacagct ggaagcctgt 360
caaagagaga gagcactttc cgtgaagttt gaggtccagg ctccacgctg gggcaacccc 420
cgtgcgctca gcttcgtact gagttcgctc cagctcgggg aggggtgga gttcgatgtg 480
ctgcctgcct ttgatgcctt gggtcagttg actggcagct ataaacctaa ccccaaatc 540
tatgtcaagc tcatcgagga gtgcaccgac ctgcagaaag agggcgagtt ctccacctgc 600
ttcacagaac tacagagaga ctctctgaag cagcgcacca ccaagctcaa gagcctcatc 660

```

```

cgcttagtca agcactggta ccaaaattgt aagaagaagc ttgggaagct gccacctcag 720
tatgccctgg agctcctgac ggtctatgct tgggagcgag ggagcatgaa aacacatttc 780
aacacagccc aaggatttcg gacgggtcttg gaattagtca taaactacca gcaactctgc 840
atctactgga caaagtatta tgactttaaa aacccccatta ttgaaaagta cctgagaagg 900
cagctcacga aacccaggcc tgtgatacct gacccggcgg accctacagg aaacttgggt 960
ggtggagacc caaaggggtg gaggcagctg gcacaagagg ctgaggcctg gctgaattac 1020
ccatgcttta agaattggga tgggtcccca gtgagctcct ggattctgct ggctgaaagc 1080
aacagtacag acgatgagac cgacgatccc aggcagctatc agaaatatgg ttacattgga 1140
acacatgagt accctcattt ctctcataga cccagcacgc tccaggcagc atccacccca 1200
caggcagaag aggactggac ctgcaccatc ctctgaatgc cagtgcattt tgggggaaag 1260
ggctccagtg ttatctggac cagttccttc attttcaggt gggactcttg atccagagaa 1320
gacaaagctc ctcatgagc tgggtgtataa tccaagacag aaccaagtc tcctgactcc 1380
tggccttcta tgccctctat cctatcatag ataacattct ccacagcctc acttcattcc 1440
acctattctc tgaaaatatt ccctgagaga gaacagagag atttagataa gagaatgaaa 1500
ttccagcctt gactttcttc tgtgcacctg atgggagggt aatgtctaag gtattatcaa 1560
taacaataaa aataaagcaa ataccaaaaa 1590

```

<210> 256
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_016816

<400> 256
 cgatcccagg acgtatcaga aatatgggta cattggaaca catgagtacc ctcatctctc 60

<210> 257
 <211> 2905
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_016817

<400> 257
 cggcagccag ctgagagcaa tgggaaatgg ggagtcccag ctgtcctcgg tgectgctca 60
 gaagctgggt tggtttatcc aggaatacct gaagccctac gaagaatgtc agacactgat 120
 cgacgagatg gtgaacacca tctgtgacgt ctgcaggaaac cccgaacagt tccccctggt 180
 gcagggagtg gccatagggtg gctcctatgg acggaaaaa gtcttaagag gcaactccga 240
 tggtagcctt gtccttttct tcagtgaactt aaaaacattc caggatcaga agagaagcca 300
 acgtgacatc ctcgataaaa ctggggataa gctgaagttc tgtctgttca cgaagtgggt 360
 gaaaaacaat ttcgagatcc agaagtcctt tgatgggtcc accatccagg tgttcacaaa 420
 aatcagaga atctctttcg aggtgctggc cgccttcaac gctctgagct taaatgataa 480
 tcccagcccc tggatctatc gagagctcaa aagatccttg gataagacaa atgccagtcc 540
 tggtagattt gcagtctgct tcaactgaact ccagcagaag ttttttgaca accgtcctgg 600
 aaaactaaag gatattgatcc tcttgataaa gcactggcat caacagtgcc agaaaaaaat 660
 caaggattta ccctcgctgt ctccgtatgc cctggagctg cttacgggtg atgcctggga 720
 acaggggtgc agaaaagaca actttgacat tgctgaaggc gtcagaacgg ttctggagct 780
 gatcaaatgc caggagaagc tgtgtatcta ttggatggtc aactacaact ttgaagatga 840
 gaccatcagg aacatcctgc tgcaccagct ccaatcagcg aggccagtaa tcttgatcc 900
 agttgaccca accaataatg tgagtggaga taaaatatgc tggcaatggc tgaaaaaaga 960
 agctcaaacct tgggtgactt ctccaacct ggataatgag ttacctgcac catcttgaa 1020
 tgctctgcct gcaccactct tcacgacccc aggccacctt ctggataagt tcatcaagga 1080
 gtttctccag cccaacaaat gcttccatga gcagattgac agtgctgtta acatcatccg 1140

tacattcctt aaagaaaact gcttccgaca atcaacagcc aagatccaga ttgtccgggg 1200
 aggatcaacc gccaaaggca cagctctgaa gactggctct gatgccgac tcgtcgtgtt 1260
 ccataactca cttaaaagct acacctccca aaaaaacgag cggcacaaaa tcgtcaagga 1320
 aatccatgaa cagctgaaag ctttttgag ggagaaggag gaggagcttg aagtcagctt 1380

tgagcctccc	aagtggaagg	ctcccagggg	gctgagcttc	tctctgaaat	ccaaagtcct	1440
caacgaaagt	gtcagctttg	atgtgcttcc	tgcttttaat	gcactgggtc	agctgagttc	1500
tggctccaca	cccagccccg	aggtttatgc	agggtcattt	gatctgtata	aatcctcgga	1560
cctcccggga	ggagagtttt	ctacctgttt	cacagtcctg	cagcgaaact	tcattcgctc	1620
ccggcccacc	aaactaaagg	atttaattcg	cctgggtgaag	cactggtaca	aagagtgtga	1680
aaggaaactg	aagccaaaagg	ggtctttgcc	cccaaagtat	gccttggagc	tgctcaccat	1740
ctatgcctgg	gagcagggga	gtggagtgcc	ggattttgac	actgcagaag	gtttccggac	1800
agtccctggag	ctgggtcacac	aatatcagca	gctcggcatc	ttctggaagg	tcaattacaa	1860
ctttgaagat	gagaccgtga	ggaagtttct	actgagccag	ttgcagaaaa	ccaggcctgt	1920
gatcttggac	ccaggcgaac	ccacaggtga	cgtgggtgga	ggggaccgtt	ggtgttggca	1980
tcttctggac	aaagaagcaa	aggttaggtt	atcctctccc	tgcttcaagg	atgggactgg	2040
aaacccaata	ccaccttgga	aagtgcgcac	aatgcagaca	ccaggaagtt	gtggagctag	2100
gatccatcct	attgtcaatg	agatgttctc	atccagaagc	catagaatcc	tgaataataa	2160
ttctaaaaga	aacttctgga	gatcatctgg	caatcgcttt	taaagactcg	gctcaccgtg	2220
agaaagagtc	actcacatcc	attcttccct	tgatggctcc	tattcctcct	tcccttgccct	2280
tcttggactt	cttgaaatca	atcaagactg	caaacctttt	cataaagctg	ccttgctgaa	2340
ctcctctctg	caggagccct	gcttaaaata	gttgatgtca	tcactttatg	tgcatcttat	2400
ttctgtcaac	ttgtatTTTT	ttttcttgta	tttttccaat	tagctcctcc	tttttccctc	2460
cagtctaaaa	aaggaatcct	ctgtgtcttc	aaagcaaagc	tctttacttt	ccccttggtt	2520
ctcataactc	tgtgatcttg	ctctcgggtg	ttccaactca	tccacgtcct	gtctgtttcc	2580
tctgtataca	aaaccctttc	tgcccctgct	gacacagaca	tcctctatgc	cagcagccag	2640
gccaaccctt	tcattagaac	ttcaagctct	ccaaaggctc	agattataac	tggtgtcata	2700
tttatatgag	gctgttgtct	tttcttctct	agcctgcctt	tatcccccca	cccaggagta	2760
tcctcttgcc	aaagcaaaaag	actttttcct	tggcttttagc	cttaaagata	cttgaaggtc	2820
taggtgcttt	aacctcacat	accctcactt	aaacttttat	cactgttgca	tataaccagtt	2880
gtgatacaat	aaagaatgta	tctgg	2905			

<210> 258

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016817

<400> 258

aaggcttagg tgctttaacc tcacataccc tcaacttaaac ttttatcact gttgcatata 60

<210> 259

<211> 2054

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017414

<400> 259

gggaagctcg	ggccggcagc	gtttccccgc	acgctggcgc	ccagctcccc	gcgcggaggc	60
cgtgtgaagt	ttcgctttcc	attcagtgga	aaacgaaagc	tgggcggggg	gccacgagcg	120
cgggggcaga	ccaaggcggg	cccggagcgg	aacttcgggtc	ccagctcggt	ccccggctca	180
gtcccagcgt	ggaactcagc	agcggaggct	ggacgcttgc	atggcgcttg	agagattcca	240
tcgtgcctgg	ctcacataag	cgcttccctg	aagtgaagtc	gtgctgtcct	gaacgcgggc	300
caggcagctg	cggcctgggg	gttttgaggt	gatcacgaat	gagcaaggcg	tttgggctcc	360
tgaggcaaat	ctgtcagtc	atcctggctg	agtcctcgca	gtccccggca	gatcttgaag	420
aaaagaagga	agaagacagc	aacatgaaga	gagagcagcc	cagagagcgt	cccagggcct	480
gggactaccc	tcatggcctg	gttgggtttac	acaacatttg	acagacctgc	tgctttaact	540
ccttgattca	ggtgttcgta	atgaatgtgg	acttcaccag	gatattgaag	aggatcacgg	600
tgcccagggg	agctgacgag	cagaggagaa	gcgtcccttt	ccagatgctt	ctgctgctgg	660
agaagatgca	ggacagcccg	cagaaagcag	tgcggccctt	ggagctggcc	tactgcctgc	720
agaagtgcaa	cgtgcccttg	tttgtccaac	atgatgctgc	ccaactgtac	ctcaaaactc	780
ggaacctgat	taaggaccag	atcactgatg	tgcacttggt	ggagagactg	caggccctgt	840

```

atacgatccg ggtgaaggac tccttgattt gcgttgactg tgccatggag agtagcagaa 900
acagcagcat gctcaccctc ccactttctc tttttgatgt ggactcaaag cccctgaaga 960
cactggagga cgccctgcac tgcttcttcc agcccaggga gttatcaagc aaaagcaagt 1020
gcttctgtga gaactgtggg aagaagaccc gtgggaaaca ggtcttgaag ctgaccatt 1080
tgccccagac cctgacaatc cacctcatgc gattctccat caggaattca cagacgagaa 1140
agatctgcca ctccctgtac ttccccaga gcttgattt cagccagatc cttccaatga 1200
agcgagagtc ttgtgatgct gaggagcagt ctggagggca gtatgagctt tttgctgtga 1260
ttgcgacagt ggggaatggca gactcggtc attactgtgt ctacatccgg aatgctgtgg 1320
atggaaaatg gttctgtctc aatgactcca atatttgctt ggtgtcctgg gaagacatcc 1380
agtgtacctc cggaaatcct aactaccact ggcaggaaac tgcatactt ctggtttaca 1440
tgaagatgga gtgctaattg aaatgcccaa aaccttcaga gattgacacg ctgtcatttt 1500
ccatttccgt tcctggatct acggagtcct ctaagagatt ttgcaatgag gagaagcatt 1560
gttttcaaac tatataactg agccttattt ataattaggg atattatcaa aatatgtaac 1620
catgaggccc ctccaggtcct gatcagtcag aatggatgct ttcaccagca gaccggcca 1680
tgtggctgct cggctcctggg tgctcgtgac tgtgcaagac attagccctt tagttatgag 1740
cctgtgggaa cttcaggggt tcccagtggt gagagcagtg gcagtgggag gcatctggg 1800
gcaaagggtc agtggcaggg ggtatttcag tattatacaa ctgctgtgac cagacttgta 1860
tactggctga atatcagtcg tgtttgtaat ttttcacttt gagaaccaac attaatcca 1920
tatgaatcaa gtgttttgta actgctattc atttattcag caaatattta ttgatcatct 1980
cttctccata agatagtgtg ataaacacag tcatgaataa agttattttc cacaaaaaaa 2040
aaaaaaaaa aaaa 2054

```

<210> 260

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017414

<400> 260

```

tgagcatctc ttctccataa gatagtgtga taaacacggg catgaataaa gttattttcc 60

```

<210> 261

<211> 3638

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017523

<400> 261

```

ggtagatgag gctgtgacag cagcaaagaa tgacggccaa gggcgacagc aggggctggc 60
catgctgtaa aggggcttct tgggagggtc cagcctcagg aatcaagggg aactcctgag 120
ccgagaattc tgaagatctc ctccctccct gaagctgtgg gctgggccat cggaaaactt 180
tcagttttgt ttcttgcct gcaagaaacg aaactcaacc gaaagcctgc agagagcaga 240
acatggaagg agacttctcg gtgtgcagga actgtaaaag acatgtagtc tctgccaact 300
tcaccctcca tgaggcttac tgctgcgggt tcctggtcct gtgtccggag tgtgaggagc 360
ctgtcccca ggaaccatg gaggagcact gcaagcttga gcaccagcag gttgggtgta 420
cgatgtgtca gcagagcatg cagaagtctc cgctggagtt tcataaggcc aatgagtgcc 480
aggagcggcc tgttgagtgt aagtctgca aactggacat gcagctcagc aagctggagc 540
tccacgagtc ctactgtggc agccggacag agctctgcca aggctgtggc cagttcatca 600
tgcaccgcat gctcggccag cacagagatg tctgtcgcag tgaacaggcc cagctcggga 660
aaggggaaag aatttcagct cctgaaaggg aaatctactg tcattattgc aaccaaatga 720
ttccagaaaa taagtatttc caccatatgg gtaaatgttg tccagactca gagtttaaga 780
aacactttcc tgttgaaaat ccagaaatc ttcttcatc tcttccaagt caagctgctg 840
aaaaatcaaac ttccacgatg gagaaagatg ttctgcaaaa gacaagaagt ataaacagat 900
ttctctttca ttctgaaagt tcatcaaaga aagcaccacg aagcaaaaac aaaaccttgg 960
atccactttt gatgtcagag cccaagccca ggaccagctc ccctagagga gataaagcag 1020
cctatgacat tctgaggaga tgttctcagt gtggcatcct gcttcccctg ccgatcctaa 1080
atcaacatca ggagaaatgc cgggtggttag cttcatcaaa aggaaaacaa gtgagaaatt 1140

```



```

tcagctagat ttggaaaagg aaaggtacta caaattcaaa agatttcact ttttaacactg 1200
gcattcctgc ctacttgctg tgggtggtctt gtgaaagggtg atgggtttta ttcgttgggc 1260
tttaaaagaa aaggtttggc agaactaaaa acaaaactca cgtatcatct caatagatac 1320
agaaaaggct tttgataaaa ttcaacttga cttcatgtta aaaaccctca acaaaccagg 1380
cgtcgaaggga acatacctca aaataataag agccatctat gacaaaacca cagccaacat 1440
catactgaat gagcaaaagc tggagcatta ctcttgagaa gtagaacaag gcacttcagt 1500
cctattcaac atagtactgg aagtcctcgc cacagcaatc aggcaagaga aagaaataaa 1560
aggcaaccaa aaagaaaagg agtcgaagta tctctgtttg cagacgatat gattctatat 1620
ctagaaaacc ccatgatctt ggcccaaaag ctctagatc tgataaaciaa cttcagctaa 1680
ctttcaggag acaaaatcaa tatacaaaat atggtagcat ttttatacac caacgacatc 1740
caagctgaga gccaaatcaa gaatgcaatc ctattcaciaa ttgccacaaa aagaataaaa 1800
tacctaggaa tacagctaac cagggagatg aaagatctct acaacaaaaa ttacaaaaca 1860
ctgctgaaaag aaatcagaga tgacacaaat ggaaaaacat tccatactta tggataggaa 1920
gaatcaatat tgttaaaatg gccatactac ccaaagcaat ttatagattc aatgctattc 1980
ctatcaaact accaataaca ttcttcacag aatcagaaaa aaaaagcatt aaaatttatt 2040
tgaaaccaaa aaagagccca aaaagccaaa gcaatcctaa gcaaaaagaa caaagctgga 2100
ggcatcgcat tacccaactt caaactatac tacagggtca cagtaaccaa aactgcatga 2160
tactggtaca aaagcatggt gctggtacaa aagcagacac atagatcaat ggaacagaat 2220
agaggggcca gaaataaagc tacacaccta caacctcta atctttgaca aagttgacaa 2280
aaatacgcaa tggggaaaga attccccatt cagtaagtgg tactgggata actagctagc 2340
catatgcaga ggattgaaac tgaaccactt ccttacacca tatgcaaaaa tcaactcaag 2400
atggattaaa gacttaaatg taaaacccca aactataaaa actctggaag ataacctagg 2460
caataccatt ctggacatag gaacggaaaa agatttcatg acaaaagatc caaaaaataa 2520
tgtaacgaaa gcaaaaattg acaaatggga catgattaaa cagaattacc atttgactca 2580
gcaatcccat tattggttat atacccaaag gaatctaaat cattctgtca taaagacata 2640
tatacacaaa tgttcacggc agcactatac acaatcgcaa agtcaggga tcaaactaaa 2700
tgtccatcag tggtagaaaag gataaagaaa atgtggtggc agggagtggg ggctcatgtc 2760
tgtaatccca gcactttggg aggtgaggc gggtggttca cctgaggtca ggagtttgag 2820
accagcctgg ccaacatggc gaaactccgt ctccgctaaa aatacgaaaa ttagccaggc 2880
gtggtggcga gcacctgtca toccagctac ttgggaggcc taggcgtgag aatcgcttga 2940
aocctggaag tgggtggttc agtgagccga gatcctgcca ctgactcca gcctgggcaa 3000
ccaagcgaga ctctgcctta aaaaaaaaaa aaagaaaatg tggcacatat acaccatgga 3060
atactatgca gccataaaaa agaattggat catgtcctgt gcagcaacgt ggatggagct 3120
ggaagccatt atcctaaatg aactcactca gaaacagaaa accaaatacc acatgttctc 3180
acttataagt agaagctaaa cattgagtac acatggatac aaagaaggga accgcagaca 3240
ctggggccta cctgaggtcg gagcatggaa ggagggtgag gatcaaaaaa ctacctatct 3300
ggtcatatgc tttttatctg gatgatgaaa taatctgtac acaaaaccct ggtgacatgc 3360
aatttaccta tatagcaagc ctacacatgt gccctgaac ctaaaaaaa agttaaaaga 3420
aaaacgtttg gattattttc cctctttcga acaaagacat tggtttgccc aaggactaca 3480
aataaaccaa cgggaaaaaa gaaaggttcc agttttgtct gaaaattctg attaaacctc 3540
tggggccctac agcctggaga acctggagaa tcctacaccc acagaaccog gctttgtccc 3600
caaagaataa aaacacctct ctaaaaaaa aaaaaaaa 3638

```

<210> 262

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_017523

<400> 262

ttggaaaagg aaaggtacta caaattcaaa agatttcact ttttaacactg gcattcctgc 60

<210> 263

<211> 2461

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018410

<400> 263

```
atgctgggta cgctgcgcgc catggagggc gaggacgtgg aagacgacca gctgctgcag 60
aagctcaggg ccagtcgccg ccgcttcag aggcgcagtc agcggctgat agagaagtac 120
aaccagccct tcgaggacac cccgggtgtg caaatggcca cgctgacctc cgagacgcca 180
cagggattga gaatttgggg tggaagacta ataaaggaaa gaaacaaagg agagatccag 240
gactcctcca tgaagcccgc ggacaggaca gatggctccg tgcaagctgc agcctggggg 300
cctgagcttc cctcgcccg cagctcctg ggagccgatt caaaaagcgg tagggtcgat 360
gccacgtcag accaggaaga gtcagttgct tgggccttag cacctgcagt gcctcaaagc 420
cctttgaaaa atgaattaa aaggaaatac ttgacccaag tggatatact gctacaaggt 480
gcagagtatt ttgagtgtgc aggtaacaga gctggaaggg atgtacgtgt gactccgctg 540
ccttcactgg cctcacctgc cgtgcctgcc cccggatact gcagtcgtat ctccggaaa 600
agtcttggtg acccagcgaa accagcttca tctcccagag aatgggatcc tttgcatcct 660
tctccacag acatggcctt agtacctaga aatgacagcc tctccctaca agagaccagt 720
agcagagctt tcttaagcag ccagcccttt acatttgcaa tgtgaccatc 780
agtgacctgt acgcagggat gctgcactcc atgagccggc tgttgagcac aaagccatca 840
agcatcatct ccacaaaaac gttcatcatg caaaaactgga actgcaggag gaggcacaga 900
tataagagca ggatgaacaa aacatattgc aaaggagcca gacgttctca gaggagctcc 960
aaggagaact tcataccctg ctctgagcct gtgaaagggg caggggcatt aagagattgc 1020
aagaacgtat tagatgtttc ttgccgtaag acaggtttta aattggaaaa agcttttctt 1080
gaagtcaaca gaccccaaat ccataagtta gatccaagtt ggaaggagcg caaagtgaca 1140
ccctgaagat attcttcctt gatttacttc gactccagtg caacatataa tcttgatgag 1200
gaaaatagat ttaggacatt aaaatggtta atttctcctg taaaaatagt ttccagacca 1260
acaatacgac agggccatgg agagaaccgt cagagggaga ttgaaatccg atttgatcag 1320
cttcacgagg aatattgcct gagtcccagg aaccagcctc gccgatgtg cctcccgac 1380
tcctgggcca tgaacatgta cagagggggg cctgcgagtc ctggtggcct tcagggctta 1440
gaaacccgca ggctgagttt accttccagc aaagcaaaaag caaaaagttt aagttaggct 1500
tttgaaaacc taggcaaaaag atctctggaa gcaggtaggt gcctgccccaa gagcgattca 1560
tcttcacac ttccaaagac caaccccaca cacagcgcaa ctgcccgcga gcagacatct 1620
gaccttcacg ttcagggaag tagttctgga atatttagaa agtcagtgtc acccagcaaa 1680
actctttcag tcccagataa agaagtgccg ggccacggaa ggaatcgtaa cgatgaaatt 1740
aaagaagaat ttgacaagct tcatcaaaaag tattgcctca aatctcctgg gcagatgaca 1800
gtgcctttat gtattggagt gtctacagat aaagcaagta tggaagtto atatacaaca 1860
gaaggcttct taggaaaatt aaatccagac cctcacttcc agggtttcca gaagttgcc 1920
tcatacacc cttgggtgcag aaaaagtcta ctgggctcaa ctgcaattga ggctccttca 1980
tctacatgtg ttgctcgtgc catcacgagg gatggcacga gggaccatca gttccctgca 2040
aaaagaccca ggctatcaga accccagggc tccggacgcc agggcaattc cctgggtgcc 2100
tcagatgggg tggacaacac cgtcagaccg ggagaccagg gcagctcttc acagcccaac 2160
tcagaagaga gaggagagaa cacgtcttac aggatggaag agaaaagtga tttcatgcta 2220
gaaaaattgg aaactaaaag tgtgtagcta gggtatttcg gagtgttatt tatcttcca 2280
cttgctctct gtttgtatct ttgttttgtt ttgattctt gagactgtga ggacttggt 2340
gaattctctg cccttaaagt aaatattagt gaaattggtt ccatcagaga taacctcgag 2400
ttcttggtgt agaaattatg tgaataaagt tgctcaatta gaaaaaaaaa aaaaaaaaaa 2460
```

a 2461

<210> 264

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018410

<400> 264

agtgatttca tgctagaaaa attggaaact aaaagtgtgt agctaggtta tttcggagtg 60

<210> 265

<211> 1405

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018455

<400> 265

```

cacctcgctc gcagcctccc cagcgcagca gcccggtgt gggcctgcgg cagccgggtc 60
ttcctggtcc ccacctcctg gggccgacgg gcggcaggaa ggggctcggc gggacgcgcc 120
gtcaggggacc tgaggaggaa caacggaacg cgttcggaac ggctggact cccgagactc 180
acccgactcg tggccacacc gggagaactg aagcggcagt agccggcgga gacgcccagc 240
ccgaaggccg gctgctaggg agcagacagc tgaaccgctt gccagacgcc gaaacccagt 300
gacgcctcc accgctccac cgtgctccc gctccccgcc cccgcgcgcc gcgggccccca 360
aggcgcatgc gccgcctgtc ctggaggggc ccatttcctg ccgtcgtggg gggaggcaca 420
gtgagtcac tggggcacgg cagcgtctaa gccacaagcc gagcacataa gccaggctcct 480
aacggagcct atgtgtaagt ccactactgg tgcaagggtg cacacttcta agaagagcgg 540
cgtggggggc tcggcgacct tcgcttcagt cgctcccccg tgcagtcccc tgtgccaag 600
acacagcctg atgcttgtgc tccggtgggc ggagcttgga gccggcgga actgcaattg 660
gtggctttga aggcgcggcg agcgggaaca gctcttgagg agtgagactg caggagatgt 720
gggcccgtgc aaagagatgg atgagactgt tgctgagttc atcaagagga ccatcttgaa 780
aatccccatg aatgaactga caacaatcct gaaggcctgg gatTTTTTgt ctgaaaatca 840
actgcagact gtaaatTTcc gacagagaaa ggaatctgta gttcagcact tgatccatct 900
gtgtgaggaa aagcgtgcaa gtatcagtga tgctgccctg ttagacatca tttatatgca 960
atttcacag caccagaaag tttgggatgt ttttcagatg agtaaaggac caggtgaaga 1020
tgttgacctt tttgatatga aacaatttaa aaattcgttc aagaaaattc ttcagagagc 1080
attaaaaaat gtgacagtca gcttcagaga aactgaggag aatgcagtct ggattcgaat 1140
tgccctggga acacagtaca caaagccaaa ccagtacaaa cctacctacg tgggtgacta 1200
ctcccagact ccgtacgcct tcacgtcctc ctccatgctg aggcgcaata caccgcttct 1260
gggtcaggag ttagaagcta ctgggaaaat ctacctccga caagaggaga tcattttaga 1320
tattaccgaa atgaagaaag cttgcaatta gtgaacatga aaggaaaata aaaattcctc 1380
acagtcacaaa aaaaaaaaaa aaaaa 1405

```

<210> 266

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018455

<400> 266

```

ccgacaagag gagatcattt tagatattac cgaaatgaag aaagcttgca attagtgaac 60

```

<210> 267

<211> 927

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018465

<400> 267

```

ggcagcgggc gaaaggagcc ggggcctgga ggtttgcgta ccggtgcgct ggtcccgga 60
ccagcgcgcc ccagtgtggt ttcccataag gaagctcttc ttctgcttg gcttccacct 120
ttaacccttc cacctgggag cgctctctaa cacattcaga ctacaagtc agaccagga 180
gagcaaggcc cagaaagagg tcaaaatggg gtttatattt tcaaatcta tgaatgaaag 240
catgaaaaat caaaaggagt tcatgcttat gaatgctcga cttcagctgg aaaggcagct 300
catcatgcag agtgaaatga gggaaagaca aatggccatg cggattgcgt ggtctcgga 360
attcctcaaa tattttggaa ctttttttgg ccttgagacc atctctttaa cagctggagc 420
gattaaaaaa aagaagccag ctttcctggt ccgattggtt ccattaagct ttatcctcac 480
ctaccagtat gacttgggct atggaaccct tttagaaaga atgaaagggt aagctgagga 540
catactggaa acagaaaaga gtaaatgca gctgccaaga ggaatgatca cttttgaaag 600
cattgaaaaa gccagaaagg aacagagtag attcttcata gacaaatgaa atcatgctta 660

```

```

ccaatcaa at ctcaaagcac agaattattg acttgaatca tggtttttac agtttttttaa 720
atgctcaaga ttttgatatt atagatttta ttttaaaata ttaaaatgca agatagtttt 780
gagctatttt aaaataaaat ttataacatt caacacaaaa tcatggaggt gctctaaata 840
acttttagat ttcctctctc tgtgtgcatt accaatatct aagtgtaaaa ttaataaatt 900
gttttgaatt cctggaaaaa aaaaaa 927

```

<210> 268

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018465

<400> 268

```

ggaacagagt agattcttca tagacaaatg aaatcatgct taccaatcaa atctcaaagc 60

```

<210> 269

<211> 1047

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018487

<400> 269

```

cccacttctc cagccagcgc cccagccctc ccgccgcccg ctgcaggtc cccaggagcg 60
cagactgtgt ccctgacaat gggaacagcc gacagtgtg agatggcccc ggaggcccca 120
cagcacaccc acatcgatgt gcacatccac caggagtctg ccctggccaa gctcctgctc 180
acctgctgct ctgcgtgctg gccccgggccc acccaggcca ggggcagcag ccggtgctg 240
gtggcctcgt gggatgagca gatcgtgctg gggatcttga gtgcagtcct aggaggattt 300
ttctacatcc gcgactacac cctcctcgtc acctcgggag ctgccatctg gacaggggct 360
gtggctgtgc tggctggagc tgctgccttc atttacgaga aacgggggtg tacatactgg 420
gccctgctga ggactctgct aacgctggca gctttctcca cagccatcgc tggcctcaaa 480
ctttggaatg aagatttccg atatggctac tcttattaca acagtgcctg ccgcatctcc 540
agctcgagtg actggaacac tccagcccc actcagagtc cagaagaagt cagaaggcta 600
cacctatgta cctccttcat ggacatgctg aaggccttgt tcagaaccct tcaggccatg 660
ctcttgggtg tctggttctt gctgcttctg gcatctctga cccctctgtg gctgtactgc 720
tggaagaatg tcccaaccaa agggaaaaga gaccagaagg aaatggttga agtgagtgga 780
atctagccat gcctctcctg attattagt cctgggtgct ctgcaccggg cgtccctgca 840
tctgactgct ggaagaagaa ccagactgag gaaaagaggc tcttcaacag cccagttat 900
cctggcccca tgaccgtggc cacagccctg ctccagcagc acttgcccat tccttacacc 960
ccttccccat cctgctccgc ttcatgtccc ctccctgagta gtcagtgtat aataaactct 1020
catgttattg ttcccaggaa aaaaaa 1047

```

<210> 270

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018487

<400> 270

```

aaccaaaggg aaaagagacc agaaggaaat gttggaagtg agtggaatct agccatgctt 60

```

<210> 271

<211> 2280

<212> DNA

<213> Homo sapiens

<300>

<308> U17077

<400> 271

```
ccgcccgcca ccagctacgc cccgtccgac gtgccctcgg gggtcgcgct gttccctcacc 60
atccctttcg ccttcttcct gcccgagctg atatttgggt tcttggctctg gaccatggta 120
gccgccaccc acatagtata ccccttgctg caaggatggg tgatgtatgt ctgcgtcacc 180
tcgtttctca tctccttgat gttcctgttg tcttacttgt ttggatttta caaaagattt 240
gaatcctgga gagtcttggg cagcctgtac cacgggacca ctggcatcct gtacatgagc 300
gctgccgtcc tacaagtaca tgccacgatt gtttctgaga aactgctgga cccaagaatt 360
tactacatta attcggcagc ctcgttcttc gccttcatcg ccacgctgct ctacattctc 420
catgccttca gcatctatta ccaactgatgc acaggcgcca ggccaagggg gaaatgctct 480
ttgaaagctc caattattgg tccccaaaag cagcttccaa cgtttgccat ctggatgaca 540
aacggaagat ccactaaaac gtccacggga ttaacagaac gtccttgca gactgagcag 600
gacaccacac ttgttttggg catttaaat cactctgctg aataggagga agcttttctt 660
tttctcggga aaacaactgt ctcttggaat tatctgacca tgaacttgct cttctagaca 720
actcacatca aagccctcac tccactaatg gagaatccta gcccactaa tgccaagtct 780
gtttggggat ttgcctcag ctatgggctt ccctagagta ggtctagggg aatactcagt 840
ctgatctttt tttgtttgt tttattttgt tttttttgag acggagtctc gctcttcctc 900
caaggctgga gtgcagtga gcatctcca cgcactgcag gctccgcctc ccgggttccc 960
gccattctcc tgcctcagcc tcccgagtag ccgggactac aggcgccacc caccatggcc 1020
ggctaattta gttgtatttt tagtagagat ggggtttcac cgtattagcc aggatggtct 1080
cgatctcctg acctcgtgat ccgccgcct cggcctccca aagtgcctggg attacaggcg 1140
tgagccaccg tgcccgccct gattctctta aaattgaaga ggtgctgcca aggccttcag 1200
atctaacgca gatgcataga cctgttcct ggtacttgtt cagcctgtgc tggggagccg 1260
tgggtcccgag ttccctggga ggctgacagg gtcaagccac cctgcccacc accctcccac 1320
ttccctccc ctttctctc cagcattagg attcaaggga aatctgcatg aagccaattt 1380
tgagggtaga cgtgtgggga aaataaatca ttatacagta agacctggg cttgaggggt 1440
ggggaatggg gagggaagg catagcctgc tctccatga gtctgacatc tcggaaactg 1500
agcagctgcc ggacgcctgg gtcaggaatc caagacccca cctcttaagg actggttcct 1560
cagaaagcac cctcagggaa aaaggtgaaa acattacatc cgtggattct cctgccacaa 1620
ccgatttga agaaaaggct gccgcaacat ctacgcagg agtgaaggac catgtccca 1680
ggaaccgcgc tgcgccacct gcactcacc cctcacatt ctcttaagca ccgggtggcc 1740
ctcccgaggc ggcggaatgg tgggtcccac ggggttggg aagggtcac caggacctca 1800
acgggcaaa ttgtgcacac taaaatatca aatcaagggt cttggttta agtaaatgt 1860
ttttctaaag aaagctgtgt tcttctgttg acccagacga atagggcaca gccctgtaac 1920
tgacgtgcc ttctgtcatt gggaatgaaa taaattatta cgagaaagg acttgtccta 1980
actggtttga gcccttacag tttgtatct acatttttcc cctcctggg tttgcgggga 2040
cagggacaga actacaggag tcatgggaaa gaaaattctg gcttactac tgctcactgc 2100
tcactttctg atcactctga tactttttt tttttttt ttttgcaacc tgataccttg 2160
aaaagcttct atgtgtctct cctttgttg cctggcagct gtctaggatg atcactgatt 2220
actatttact aagtagccac atgcaaataa aagttgtttg gtaaatgga aaaaaaaaa 2280
```

<210> 272

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> U17077

<400> 272

```
tcaccaggac ctcaacgggc aaagttgtgc aactaaaaat atcaaatcaa ggtgcttggg 60
```

<210> 273

<211> 2554

<212> DNA

<213> Homo sapiens

<300>

<308> X87949

<400> 273

```

agggtcgacgc cggccaagac agcacagaca gattgaccta ttgggggtgtt tcgcgagtg 60
gagaggggaag cgccgcggcc tgtattttcta gacctgcctt tcgcctgggt cgtggcgcc 120
tgtgaccccg ggcccctgcc gcctgcaagt cggaaattgc gctgtgctcc tgtgctacgg 180
cctgtgggtg gactgcctgc tgctgcccac ctggctggca agatgaagct ctccctgtg 240
gcccgatgc tgctgtaagt cagcgccggc cgggcccagg aggaggacaa gaaggaggac 300
gtgggcacgg tggtcggcat cgacttgggg accacctact cctgcgtcgg cgtgttcaag 360
aacggccgcg tggagatcat cgccaacgat caggggcaacc gcatcacgcc gtcctatgtc 420
gccttcactc ctgaagggga acgtctgatt ggcgatgcg ccaagaacca gctcacctcc 480
aaccocgaga acacggtctt tgacgccaag cggctcatcg gccgcacgtg gaatgacccg 540
tctgtgcagc aggacatcaa gttcttgccg ttcaagggtg ttgaaaagaa aactaaacca 600
tacattcaag ttgatattgg aggtgggcaa acaaagacat ttgctcctga agaaatttct 660
gcatgggttc tcaactaaaat gaaagaaaac gctgaggctt atttgggaaa gaaggttacc 720
catgcagttg ttactgtacc agcctatttt aatgatgcc aacgccaagc aaccaaagac 780
gctggaacta ttgtggcct aaatgttatg aggatcatca acgagcctac ggcagctgct 840
attgcttatg gcctggataa gagggagggg gagaagaaca tcctgggtgt tgacctgggt 900
ggcggaacct tcgatgtgtc tcttctcacc attgacaatg gtgtcttcga agttgtggcc 960
actaatggag ataactcatc ggggtggagaa gactttgacc agcgtgtcat ggaacacttc 1020
atcaaaactgt acaaaaagaa gacgggcaaa gatgtcagga aggacaatag agctgtgcag 1080
aaactccggc gcgaggtaga aaaggccaag gccctgtctt ctgagcatca agcaagaatt 1140
gaaattgagt ccttctatga aggagaagac ttttctgaga ccctgactcg ggccaaattt 1200
gaagagctca acatggatct gttccgggtc actatgaagc ccgtccagaa agtgttgga 1260
gattctgatt tgaagaagtc tgatattgat gaaattgttc ttgttggtgg ctgactcga 1320
attccaaaga ttcagcaact ggttaaagag ttcttcaatg gcaaggaacc atcccggtgg 1380
ataaaccag atgaagctgt agcgtatggt gctgctgtcc aggetgggtg gctctctggt 1440
gatcaagata caggtgacct ggtactgctt catgtatgtc cccttacct tggatttgaa 1500
actgtaggag gtgtcatgac caaactgatt ccaagtaata cagtgggtgc taccagaac 1560
tctcagatct tttctacagc ttctgataat caaccaactg ttacaatcaa ggtctatgaa 1620
ggtgaaagac ccctgacaaa agacaatcat cttctgggta catttgatct gactggaatt 1680
cctcctgtc ctcgtggggt ccacagatt gaagtcacct ttgagataga tgtgaatggt 1740
attcttcgag tgacagctga agacaagggt acagggaaca aaaataagat cacaatcacc 1800
aatgaccaga atcgccctgac acctgaagaa atcgaaagga tggttaatga tgcgtgagaag 1860
tttgctgagg aagacaaaaa gctcaaggag cgcattgata ctagaaatga gttggaaagc 1920
tatgcctatt ctctaaagaa tcagattgga gataaagaaa agctgggagg taaactttcc 1980
tctgaagata aggagaccat ggaaaaagct gtagaagaaa agattgaatg gctggaaagc 2040
caccaagatg ctgacattga agacttcaaa gctaagaaga aggaactgga agaaattggt 2100
caaccaatta tcagcaaaact ctatggaagt gcaggccctc ccccaactgg tgaagaggat 2160
acagcagaaa aagatgagtt gtagacactg atctgctagt gctgtaatat tgtaataact 2220
ggactcagga acttttgtta ggaaaaaatt gaaagaactt aagtctcgaa tgtaattgga 2280
atcttcacct cagagtggag ttgaactgct atagcctaag cggctgttta ctgcttttca 2340
ttagcagttg ctacatgtc tttgggtggg gggggagaag aagaattggc catcttaaaa 2400
agcgggtaaa aaacctgggt taggggtgtg gtccaccttc aaaatgttct atttaacaac 2460
tgggtcatgt gcatctggtg taggaagttt tttctacct aagtgcacc aataaatggt 2520
tgttattttac actggtcaaa aaaaaaaaaa aaaa 2554

```

<210> 274

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> X87949

<400> 274

```

aactttcctc tgaagataag gagaccatgg aaaaagctgt agaagaaaag attgaatggc 60

```

<210> 275

<211> 1359
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig1632

<400> 275
 ttttaagaca gttacctgtt gtgctgctgt tacaatatat aatgaaacca agtcagggga 60
 gtgaatttat caatcttttg atgtaaaagta aaaacgtagt tcacacttca ggagagaact 120
 tcatagcaca atgtctttct ataagatatt tttaatgatt tagtatttta caacatttgt 180
 ttaccatatt ttgatatacc atttttttct atctgcccag ttttattaaa aaaactatat 240
 attattttct aaagaaacaa tcatattttt atacaaaatt atgttttcag gtaacgaaat 300
 agatgtaggg tacagtggaa cataagcagt gttaccctct gctgggagtc agtattatac 360
 aacaaatggt gagctggaac atgocctgtc tgtgctgtcc ctccctgtgt gggtcgcgga 420
 tgtgtaggca acattgcctt atcacgctag gttcacctga cactttaaaa ggaaaaaag 480
 ttccatagag ttctgtggtc acaaaattgt tttgctttta tcaaatactt taatagaacc 540
 aaagttgcag atattggaat gtatggaagt atctcagtct ctgcataaga ggattaaagt 600
 atgaaaggat catttaatga ctgttttact tataagtcatt taagtaatcc accattttct 660
 atggatgatg cttaagcctg gtgaggtttg tactctaagg agcccagatc ataatgcagt 720
 gcatttcctt agcccttaga gtttcttgca aacattttaa aaaagacata ttttaagaaag 780
 aaagataaag aaaaaacata tttaattact gtaaacaggt actgctttat gtttattttc 840
 tctctacttc aaccaaatac agatctttga ggttttgctg acattgttgg tggttttgca 900
 catgttcttt ctaattggat ttatgaatag ttctatgggt tttcaaagat gaatcatgct 960
 aagaacactt ctgctttttg atccactgtt tgcagcagaa ttatatatat gtataggaaa 1020
 aatccacttt gaataatcca tgttttgtat ttggaaattg tttttaaaaa taaaaaggaa 1080
 aggaaatata taaagctgtt atttattctg catttcttac atatctatcg cttgtcagta 1140
 taccctgttt ggtatatatt gcctctgcac atctacattt gtatatgcaa cagtgcagctt 1200
 tatacttaca taaactgtaa ataatccttt ctgtgaaagg atcatcatat caagatgata 1260
 ccaaaagtat gtaaaaagaa acctgcatta ttttgtaatt atttcttata gatatttcat 1320
 ggtaagatta gcagtcaata aagttacttt tttgccttt 1359

<210> 276
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig1632

<400> 276
 gggttttcaa agatgaatca tgctaagaac acttctgctt ttgatccac tgtttgcagc 60

<210> 277
 <211> 994
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig3464

<400> 277
 tgaatgtata tattaagact gtagctgaat tgcacatgaa atcagattgc caacttcttg 60
 actttcaatg ttagacattt atccttaagt tgtgagcgat atatgtagca tgctgtgaaa 120
 tgtctgttat agctctttta ttcacagta ttaatacaga attatcattt gcgtttcttg 180
 gtacttttta ttcaatgtaa tcagaagctg tgatgttttg cctttgtagt cctgtgcttt 240
 gttactgtaa tttttttttt tttttttacg aagcacgtga ctggactaat gtaaggcaga 300
 tgacgtgatc tttaagactg ctatatatat cagtctctta ctctataagg ttttaaatga 360
 gaataagctt ttatcaaata gataattgat gcaatttagg attcacgcaa gtttcagtgt 420
 caaatggcgg tcttatagtt tcaattctga aaatagcaaa cttaataaac agccacttta 480
 aacttgttct ggcaaaccag acctgctgtg agatatagtc taaggtagtt aaccatataa 540

```

gccttttcaa ctcttaatgc cctccacatg aatcagcagt taagaagggt ctagaaccca 600
tgaaagcttt tgtatgtatt actagggttt gtttttctta tgtttgctga ttttacagtt 660
ctgactaaag ctgacctaaa tggatcagtt tatgtgtaat attctagtgc tttaatgact 720
ctttttttct ttggagggag ggtaacatta tttggacaga tgcagaagga actgttagtg 780
agtcaagaca aacacatctg aaataaagga actgtgtatt aacatgttaa caattcataa 840
ctgcactttt tatgacattt tgaaaatcta tttataggta cagaacaatg ggttttgtta 900
aactgtatca catttatact tgcagaaatt tatttcattg ttattagtag gaattttatt 960
ggttcaataa aattggcaaa actgaacacc aaaa 994

```

<210> 278
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig3464

<400> 278
 ctgctgtaga tatagtctaa ggtagttaac catataagcc ttttcaactc ttaatgccct 60

<210> 279
 <211> 423
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig14683

<400> 279
 tatgttatgg atatcttatt ttagagtaag aatataaggc atagccatat ttatgaaggt 60
 agtaatactc tactaatcaa tacttagaag tttttgttat gactaatctg aatgcttttt 120
 agtttttcct taatctagtt atgttggtta tttataagtc agttttcaga ttaggaaaga 180
 aggtatttga ggggtgttcca tttccactga atagtaagat gatgcttact tagatttcca 240
 cagctgtttg aaagctctgt atttggtat aacggaaaac tttgttaggg atgcttgatg 300
 ttttgtgttt tgtttctaaa ggaagacagt gttttgttcc ttcttttagaa aacttgaaga 360
 atagaataat gagtccagga ttaatttggg ataaagtctt ttacttcata aattctgatt 420

ctg 423

<210> 280
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig14683

<400> 280
 aggaagacag tgttttgttc cttctttaga aaacttgaag aatagaataa tgagtccagg 60

<210> 281
 <211> 391
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig28552

<400> 281
 atgccattga tgtgaagaag gtgtctgtgg aagactttct tactgacctg aataacttca 60
 gaaccacatt catgcaagca ataaaggaga atatcaaaaa aagagaagca gagggaaaaa 120


```

aaaaacgtgt cagaatagct aaagaattag cagagcgaga aagactcgaa cgccaacaaa 180
agaaaaagcg tttattagaa atgaagactg aggggtgatga gacaggagtg atggataatc 240
tgctggaggg cttgcagtcc ggggctgcct tccgcgacag aagaaaaagg acaccgatgc 300
caaaagatgt tcggcagagt ctccagtcct tgtctcagag gcctgttctg aaagtttgta 360
accatggtaa taaaccgtat ttataaattg c 391

```

<210> 282

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28552

<400> 282

```

aagactttct tactgacctg aataacttca gaaccacatt catgcaagca ataaaggaga 60

```

<210> 283

<211> 450

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28947

<400> 283

```

ctcatccaag gagctggggc agacttcatt gattctagag agacctgttt cagtgcctac 60
tcatccctgc cctctggtgc cagcctcctt accatcacgg cttcactgag gtgtaggtgg 120
gtttttctta aacaggagac agtctctccc ctcttacctc aacttcttgg ggtgggaatc 180
agtgatactg gagatggcta gttgctgtgt tacgggtttg agttacattt ggctataaaa 240
caatcttggt gggaaaaaat tgggggagag gacttcttcc tacacgcgca ttgagacaga 300
ttccaactgg ttaatgatat tgtttgtaag aaagagattc tgttggttga ctgcctaaag 360
agaaagggtg gatggccttc agattatacc agcttagcta gcattactaa ccaactgatg 420
gaagctctga aaataaaaga tcttgaaccc 450

```

<210> 284

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig28947

<400> 284

```

agacagattc caactgggta atgatattgt ttgtaagaaa gagattctgt tgggtgactg. 60

```

<210> 285

<211> 439

<212> DNA

<213> Homo sapiens

<300>

<308> Contig30875

<400> 285

```

agaaatcaat gacagttgac aggaagagag gacgcataca acaggcaaaa gaggaatgcc 60
cagcagtcctt ggtccttgcg gtgcaatact ggccttgagg ccaagtcagc aggggattcg 120
tagtcactaa cttctaactg aggcagggaa gtaccatggt ctggaaaagg tccaaagaaa 180
caggaataga ggcagtgtag caagaggcag atttttggtg ccaaatagat ttgaatcctg 240
gttctgcttc ttcctttgta gagtatgata ttggttcttt cctcccaaag ctattataaa 300

```

gactaaatat gtacacaaat ctttgggatg tctgacatat aaatgcttaa caataggtat 360
ttgctgggtat tattacaaat gaatttgctt atttttgagc cacttctatg tctgtccatt 420
aaaccaaaaat gtgttctgc 439

<210> 286
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig30875

<400> 286
ggttctttcc tcccaaagct attataaaga ctaaatatgt acacaaatct ttgggatgtc 60

<210> 287
<211> 338
<212> DNA
<213> Homo sapiens

<300>
<308> Contig31221

<400> 287
gggaagttac actgcttcac accacaaggc cgtgggaaat cttggagggt ctgtgccttt 60
ctgtcacctc tactttttgc agctgtgatt gcactgtccc gcacatgtga ctacaagcat 120
cactggcaag gaccctttaa atgggtgaaaa tgggcagatg aatagcaata agtggacctt 180
tgttactctt ctgagttaga aaaattctaa tttagtacac tctgaacaaa gcttattata 240
cttacttaag atgtgttttg atttgggtgt cagaaagcaa cctgacaatg ataatactgt 300
aactatgata aaattgagaa taaaaagatt ttatttag 338

<210> 288
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig31221

<400> 288
aaatgggcag atgaatagca ataagtgac ctttggttact cttctgagtt agaaaaattc 60

<210> 289
<211> 417
<212> DNA
<213> Homo sapiens

<300>
<308> Contig31288

<400> 289
gaatcacttg agcccgggag gttgaggctg cagtgaagctg tgtttatacc actgcactcc 60
agcctgctgg gtaacagagc aagactccat ctcaaaaaga aaagaaaaaa tgctttgcta 120
cataatgagg ccaggcaaaa aaaaaaaaaa tctctgtggaa atcatataga caaacatttg 180
caaagctgct actgccattg taccagtgtt aaaatgtgtt ctaccttgca tcttttactg 240
atttttatga cagattttat attgtaacca tttgagaact ctgtaagtgc tatggcttcc 300
ttaaactacg atttatcata tgctcccagt gtttactttg agactgaatg gcaaccagag 360
aatgtaaaca accaaggtgc atctggttat gttttaaaat aaagattaat aaaagtt 417

<210> 290
<211> 60

<212> DNA
<213> Homo sapiens

<300>
<308> Contig31288

<400> 290
ggcttcctta aactacgatt tatcatatgc tcccagtggt tactttgaga ctgaatggca 60

<210> 291
<211> 394
<212> DNA
<213> Homo sapiens

<300>
<308> Contig31646

<400> 291
gctgctacac cccatgtaaa aagcggaaaa taaaatgaag attttccagc gcaagatgcg 60
gtactggttg cttccacctt ttttggcaat tgtttatttc tgcaccattg tccaaggtca 120
agtggctcca cccacaaggt taagatataa tgtaatatct catgacagta tacagatttc 180
atggaaggct ccaagaggga aatttggtgg ttacaaactt cttgtgactc caacttcagg 240
tggaaaaact aaccagctga atctgcagaa cactgcaact aaagcaatta ttcaaggcct 300
tatgccagac cagaattaca cagttcaaat tattgcatac aataaagata aagaaagcaa 360
gccagctcaa ggccaattca gaattaaaga tttta 394

<210> 292
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig31646

<400> 292
gccagaccag aattacacag ttcaaattat tgcatacaat aaagataaag aaagcaagcc 60

<210> 293
<211> 357
<212> DNA
<213> Homo sapiens

<300>
<308> Contig37562

<400> 293
caattatttc aagtgcacct tattaacaaa agtatcagt gatccaacat aaaattttat 60
agtactaaat gtcaagccta actgtgaatt ttgttctgta tcttaagtaa atttatgata 120
atgttctcga gctatcaaca aaatatatgt acttttgtga gctatgaatt ttctaattaa 180
attttacatg ctataacatg atttttacat gaatgatact ttgtttataa ctatcaaatg 240
tcagtatttt actacaattt tattataaag tgtacattat cactaaatga acttcgattt 300
taaaaatcaa attagcttta gttgtatatt attttttaca aataaagata gacttgt 357

<210> 294
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig37562

<400> 294
atcaaatgtc agtatttttac tacaatttta ttataaagtg tacattatca ctaaataaac 60

<210> 295
<211> 351
<212> DNA
<213> Homo sapiens

<300>
<308> Contig37895

<400> 295
aatagagaca cctctaatta attaaagcgg atgccctccc cactcctccc aggatttgac 60
tcggagcaca aactcttcac aaacaaaaat gtcaggacac catcgccagt gtccactggc 120
cactgctgtt ggtgtgaggc agccaggagc cctcagaac tagtaagtct gagaagaggc 180
tgcacggggc ctaggagagg gagaaatgag ccggtccaag gtgaattcct tgattctcca 240
ttgtgagtgc accaagaaca agcactccct ccgactgact ctgcctacc aggatctgga 300
acaccttcca ttaattttatt cgttcattca ataaatattt attgactgac t 351

<210> 296
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig37895

<400> 296
ctctcgccta ccaggatctg gaacaccttc cattaattta ttcgttcatt caataaatat 60

<210> 297
<211> 418
<212> DNA
<213> Homo sapiens

<300>
<308> Contig38288

<400> 297
gacaagtaaa tggggggccgt tgggacggcg ggtgcctgga gggcagctct gggctcagcg 60
ggcagtgcctt agagcacagg cccctctgtt ggggggatggg gaggagagca gtctgccctt 120
gggagcgtag gccccaggga gacttctaaa gccccccctg tcgtctgctc ttcacccagc 180
accacagagg cacctgctgc acacacaagc atctcactcg gccacaggag gggggccaggc 240
ttcctttgcc tgaagctggt ttgggaaggg tctccacaca ggcactgatc tcccaagctt 300
tgggtcatgat gtcttttacc atttgataat tttaaacatt gttttttaaac ccaaaacatt 360
tagtggtccg ttgcctctga agatgtaaac aaacaaatac actatttctg ggaacatt 418

<210> 298
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig38288

<400> 298
tttagtggtc cgttgcctct gaagatgtaa acaaacaaat acactatttc tgggaacatt 60

<210> 299
<211> 413

<212> DNA
 <213> Homo sapiens

<300>
 <308> Contig38901

<400> 299
 tacatttttg tttaatgttg ggcctgaggt taactgtgac catggtccag cttgagtggc 60
 ttctggagca gccacatttt caaggactgt ccaaaagcca gccagttcag ggctcaggcc 120
 tcacccattg cccactcctg gggagaccat cacctggctc atcgtttcca ccaagagtgc 180
 cccacaggag tgccccacag acccgctgga ccagcctgct gcgggtcctg gccaggggtc 240
 tggctaacgg tgagggctga ctctgaactg tctctcagtc tccagaaagt gttcaagcct 300
 gttgtgttcc caaatctgat tcctoctatt gtcttgtaaa tcaaactcta agtgaaaact 360
 toccatttgt cccttcaaag attttttttt attaaatggt tttttaagat cct 413

<210> 300
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig38901

<400> 300
 tgttcccaaa totgattcct cctattgtct tgtaaatcaa actctaagtg aaaacttccc 60

<210> 301
 <211> 434
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig40434

<400> 301
 gaatggtgaa agagagatgc cgtgttttga aagtaagatg atgaaatgaa tttttaattc 60
 aagaaacatt cagaaacata ggaattaaaa cttagagaaa tgatctaatt tccctgttca 120
 cacaaacttt acactttaat ctgatgattg gatattttat tttagtgaat catcatcttg 180
 ttagctaaact ttaaaaaatg gatgtagaat gattaaaggt tggatgatt tttttttaat 240
 gtatcagttt gaacctagaa tattgaatta aaatgctgtc tcagtatttt aaaagcaaaa 300
 aaggaatgga ggaaaattgc atcttagacc atttttatat gcagtgtaca atttgctggg 360
 ctagaaatga gataaagatt atttattttt gttcatatct tgtacttttc tattaataatc 420
 attttatgaa atcc 434

<210> 302
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig40434

<400> 302
 aaggaatgga ggaaaattgc atcttagacc atttttatat gcagtgtaca atttgctggg 60

<210> 303
 <211> 391
 <212> DNA
 <213> Homo sapiens

<300>

<308> Contig40552

<400> 303

```
caccaagccc tgctccggca cctcgaatcc ctggcgacca tgagtcacca gctccaagcc 60
ttactgtgcc cccagaccaa gagctccatc ccccgccctc tgcagcgttt gtctagcgcc 120
cttgcaagctc cagagccccc tggcccagcc cgtgactcct ctttggggcc tacagatgaa 180
gctggctctg agtgtccctt ccctagaaag gcctgaccct ccttaccacac cagaacaggg 240
gttttgatgc cctcactagt gttgaagcct gttccagaga gaggtgggac tgcaaggaga 300
ggatggtcag ccctaccac ctgccctgtt tgagcttcct gtttgacaat gtttgctgtt 360
gattttttgt tcaataaaga atttggtaaa a 391
```

<210> 304

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig40552

<400> 304

```
tttgagcttc ctgtttgaca atgtttgctg ttgatttttt gttcaataaa gaatttggtg 60
```

<210> 305

<211> 495

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41413

<400> 305

```
aaatattctt aatagggcta ctttgaatta atctgccttt atgtttggga gaagaaagct 60
gagacattgc atgaaagatg atgagagata aatgttgatc ttttggcccc atttgtaaat 120
tgtattcagt atttgaacgt cgtcctgttt gttgttagtt ttcttcatca tttattgtat 180
agacaatttt taaatctctg taatatgata cattttccta tcttttaagt tattgttacc 240
taaagttaat ccagattata tggtccttat atgtgtacaa cattaaaatg aaaggctttg 300
tcttgcattg tgaggtacag gcggaagtgt gaatcaggtt ttaggattct gtctctcatt 360
agctgaataa tgtgaggatt aacttctgcc agctcagacc atttctaata cagttgaaag 420
ggaacaagt atttcagtct caaaattgaa taatgcacaa gtottaagt attaaaataa 480
aactgttctt atgtc 495
```

<210> 306

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41413

<400> 306

```
cagctcagac catttcctaa tcagttgaaa gggaaacaag tatttcagtc tcaaaattga 60
```

<210> 307

<211> 409

<212> DNA

<213> Homo sapiens

<300>

<308> Contig41538

<400> 307

```

aaaaaaaaa aaaaaaaaaa aaagagttgt tttctcatgt tcattatagt tcattacagt 60
tacatagttc gaaggtctta caactaatca ctggttagcaa taaatgcttc aggcccacat 120
gatgctgatt agttctcagt ttctattcag ttccacaatat aaccaccatt cctgccctcc 180
ctgccaaagg tcataaatgg tgactgccta acaacaaaat ttgcagtctc atctcatttt 240
catccagact tctggaactc aaagattaac ttttgactaa ccctggaata tctcttatct 300
cacttatagc ttcaggcatg tatttatatg tattcttgat agcaatacca taatcaatgt 360
gtattcctga tagtaatgct acaataaatc caaacatttc aactctggt 409

```

<210> 308
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig41538

```

<400> 308
ctcatgttca ttatagttca ttacagttac atagtcogaa ggtcttacia ctaatcactg 60

```

<210> 309
 <211> 552
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig41887

```

<400> 309
ctgaagacta cgaccatgaa atcacagggc tgcgggtgtc tgtaggtctt ctccctgggtga 60
aaagtgtcca ggtgaaactt ggagactcct gggacgtgaa actgggagcc ttaggtggga 120
ataccacagga agtcaccctg cagccaggcg aatacatcac aaaagtcttt gtcgccttcc 180
aagcttttct cgggggtatg gtcattgtaca ccagcaagga ccgctatttc tattttggga 240
agcttgatgg ccagatctcc tctgctacc ccagccaaga ggggcagggtg ctggtgggca 300
tctatggcca gtatcaactc cttggcatca agagcattgg ctttgaatgg aattatccac 360
tagaggagcc gaccactgag ccaccagtta atctcacata ctcagcaaac tcaccctgtg 420
gtcgttaggg tgggggtatg ggccatccga gctgaggcca tctgtgtggt ggtggctgat 480
ggtactggag taactgagtc gggacgctga atctgaatcc accaataaat aaagcttctg 540
cagaatcagt gc 552

```

<210> 310
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig41887

```

<400> 310
tactggagta actgagtcgg gacgctgaat ctgaatccac caataaataa agcttctgca 60

```

<210> 311
 <211> 745
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig42342

```

<400> 311
gcagtaaaga caggacgcac ccatgtcaca agaggagcac aggcaggggt gttggtgttg 60

```

```

gggcagccct cagggctctc agaccagcc ccactcacac agcagcctag gaaggaaggg 120
cagagtccca ggtgtcagct ggtgggtctc ccaggagctg cccctccctg gaagtcacag 180
gacaggaatg acagatcagg gaactgcagg aagctgccac ctctggggtc agaatatgcc 240
cagcctgagg gggctctcta tcgggggtct cgagagccag acagcctgcc ttgtgctgca 300
tacctggctt tgctctgtgc agaaccagc acacgtgatt ttgtgtgaca tgccagcagc 360
ctggctccca ggacaggagg cctgccctgg gggaggggct gcaggaggag ggggggcagg 420
caccatgag tctgtccagc cttgtcacag atgcatcgcc caagctgcgg tcctgatttc 480
agctcacctc agagtaaadc agaataaact gcacccagac tttcacgaat gcatgttgac 540
gctttcagtt cacccttttc tttgctaact ttcttctat tttcttctaa tgcgagagct 600
tattaattcc atatttatca ttttgaataa cttttctcct ttttagtaac aaaatgtact 660
tcactcttag taaaatgtat ttactatatt agtaacaaaa atatacttgc ctaatcatgt 720
ttaaaatata gtgatgtgaa aaatt 745

```

```

<210> 312
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig42342

```

```

<400> 312
caccagact ttcacgaatg catgttgacg ctttcagttc acccctttct ttgctaactt 60

```

```

<210> 313
<211> 398
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig43645

```

```

<400> 313
agttcaaagg cagataaatc tgtaaattat tttatcctat ctaccatttc ttaagaagac 60
attactccaa aataattaaa ttttaaggctt tatcaggctc gcatatagaa tcttaaattc 120
taataaagtt tcatgttaat gtcataggat ttttaaaaga gctataggta atttctgtat 180
aatatgtgta tattaataat taattgattt cagttgaaag tattttaaag ctgataaata 240
gcattagggt tctttgcaat gtggtatcta gctgtattat tggttttatt tactttaaac 300
atthtgaaaa gcttatactg gcagcctaga aaaacaaaca attaatgtat ctttatgtcc 360
ctggcacatg aataaaacttt gctgtgggtt actaatct 398

```

```

<210> 314
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig43645

```

```

<400> 314
gaaaagctta tactggcagc ctagaaaaac aaacaattaa tgtatcttta tgtccctggc 60

```

```

<210> 315
<211> 478
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig44289

```

```

<400> 315

```



```

ctaaaaacaa cactcatcag tcttgggaaa tttgaacttt gatcaactta actaaagaag 60
gaagggtagt aagaattttt caaatacaaa tatttgccaa ttcacagatg ataacattta 120
aggccttcaa aagtaagggt ttttccttgt tttccagtc agcttttgtc aactctaata 180
gttttttcat aaacattttt tatttgtata attgcaacag ttttaagaaat tatcacaact 240
atitagaaac atttaaaatg ttctttttga tataagctat atacttggaa aaatacattg 300
gtatctaaaa tttgagggtg gttaagactg ctttttggtt taaaaaatgg tttacattca 360
aatttttgaa gtgttttatg cttcatatgg ctaagttgta gtttggcaga gttaacagca 420
taagaataaa catgctgtaa ttttaaaaga tgctttgaat aaaaatttat ttttaattt 478

```

```

<210> 316
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig44289

```

```

<400> 316
catcagtctt gggaaatttg aactttgatc aacttaacta aagaaggaag ggtagtaaga 60

```

```

<210> 317
<211> 556
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig44909

```

```

<400> 317
accatctggg atttctacag cctgggtacc catagccaca ccaaggcttc tgggagattc 60
tgcagggtca gctttccagg ctgttcccaa atagctccct gcctccccac tgcccctaaa 120
gccacagcag aagagccatt catctcataa acaaaaagga agaggaaaaga atgaggaagg 180
accctgtgca aggttatattg caggcaggga tgggcttgta cctgacagca cccaccctg 240
tgtggcccc aggccctcat caccctcaga cccctcctaa gcagttccct cattgctctt 300
tggtactagg tgacagcagg aagagcaggg cccatgaccg ggtggaagtt cagttttggg 360
gtctgcttca agaggggggtt ttacactctg attccaggac aagcactctg aggcgggtgg 420
gggagagaaa ccctggctct tcaccaggtt ttcacacaca tgtaaataaa acactatggt 480
agtatctaac acactcctgg atacagaaca caagtcttgg cacatatgtg atggaaataa 540
agtgttttgc aatctt 556

```

```

<210> 318
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig44909

```

```

<400> 318
tcaccaggtt ttcacacaca tgtaaataaa acactatggt agtatctaac acactcctgg 60

```

```

<210> 319
<211> 710
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig45032

```

```

<400> 319
aaagataggc ttctaagtta aggcaaatca ttcattctgt cattaaacaa atacaaacca 60

```

```

ggcacctgtc atatgccaa tgatattcaa aatggcccat gtagaccttt gtgaagtatg 120
tggcctaaca gacattaaac aaatgtctgt gaaactgaca taataaagta aggtaagtta 180
tatgtgagac attctctttt tataataatt cctgtaaagc agtacttact taggtaatga 240
tatcatactg ttttgtttta ttttttcctt aagagctaaa acgtcatcct ctcttcagtg 300
atgtggactg ggaaaatctg cagcatcaga ctatgccttt catccccag ccagatgatg 360
aaacagatac ctcttatttt gaagccagga atactgctca gcacctgacc gtatctggat 420
ttagtctgta gcacaaaaat tttcctttta gtctagcctc gtgttataga atgaacttgc 480
ataattatat actccttaat actagattga tctaaggggg aaagatcatt atttaacctt 540
gttcaatgtg cttttaatgt acgttacagc tttcacagag ttaaaaggct gaaaggaata 600
tagtcagtaa tttatcttaa cctcaaaact gtatataaat cttcaaagct tttttcatct 660
atttattttg tttattgcac tttatgaaaa ctgaagcatc aataaaatta 710

```

<210> 320

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig45032

<400> 320

```

ttaacctagt tcaatgtgct tttaatgtac gttacagctt tcacagagtt aaaaggctga 60

```

<210> 321

<211> 726

<212> DNA

<213> Homo sapiens

<300>

<308> Contig46218

<400> 321

```

atacatattg ctttagagag caggtaggtg gccatgtgtt cagcagtgtg tccttaagaa 60
aataccatct ttctaagcca ctggaatttt tactttacta tttttaacat taatggatgt 120
caggtcacat acctcaagtc tttacatctc catgtatatt ccatatata tgtttatata 180
ggcccaagtt tctccttaat tgggatctat atactaccag cacaacatca aaaacatgta 240
attgaataca tcagagctat atatgtaagg aaatgactgg tgaccccatc atcatcattg 300
ttgaattcat gttaagtaga cctctaggg gaccataagg caattgagca cataacgaaa 360
aatgatgcaa taagaatgta tgcactctct ttgccaaatg catgtgcttt tgtgtaacgt 420
ggatgtaaac agaattgcag tgctgccgaa attcttgatc ttggctaaga gagtattttt 480
ccccttgtaa ttatgactct gagataaaat tgccattttg aaatttccaa agtaacaact 540
tttttatatt tatgaataaa cttgggattg caatttctct gatctgacaa tcaataactt 600
taacaaagat ctaataaagt gtttcaagga aagttttcct aagcaaagt aatattacct 660
catttgggca tcattactct gttaattcta tatcaaagga aataaacttg ctacttgcac 720
taaattg 726

```

<210> 322

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig46218

<400> 322

```

accataaggc aattgagcac ataacgaaaa atgatgcaat aagaatgtat gcactctctt 60

```

<210> 323

<211> 580

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47096

<400> 323

```
ggtggtctct catccttgtg tgctgctctg ctaagagatg tccaaggcgg agccggggca 60
agatccttcc agactcatct gtcagagccc caagcccttt agaccagag cccaaggacc 120
atgcctttgg gacatttaga ctgcagcctt tgcttctgtg tttttggag ttttggtgac 180
ttttgtcacc tggacaçact catttgtag ccatagtggg ttcccttggg cagcaacagt 240
```

```
gcattgtacct ctggatgtca tctgaggtga gaccaccgag gccttttctc tctgtgtaca 300
gaggggagtt aggagttgct ggactggatg cattacgagg actggggaca gggtagaggg 360
acatccaggg atcagggcat gagtgggggc aaccccccg cctctgccct ggcattggtct 420
ccgcatgggc tgaggtgtag ctgattggct gccacatttc ggccatgctg gctggcgtgc 480
ccatgttgca gatattttcc cgagttcccc agaattggat gtattgaatc tcagccacat 540
gcaacactgt gtccagcatt ctttgcaata aatacttttt 580
```

<210> 324

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47096

<400> 324

```
atattttccc gagttcccca gaatggatgg tattgaatct cagccacatg caacactgtg 60
```

<210> 325

<211> 632

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47563

<400> 325

```
gccatctagt ctgtgggttt ctgttgaagc agtctgaatt gactaaaaca gtcacttgga 60
gtagtataaa accactttcc tgttgaaagc agaacatgct gattcaactg ttttgttcaa 120
tagcaatgat agattttggt taagtcccct acactttctt atttctaaat gatcaagagt 180
acacttcctg gcagtgatta aggagtgtgt atctaacaga aaaaatatat ataccctgtg 240
aaccogaata tggaattcag attgtttctg ccctcagtat catacttaaa aaacaagcat 300
acaacaaaac ataagggaac aaacagcaac cataacaaaa aaaaaccttt aaagggtggg 360
ttttgctgtg ataaatgaat acggtactct gaaggagaaa aaagtttctc aaatgagctt 420
aaactgcaag tgatttaaaa attagagaat ataattctta aagctattga aagtttcaac 480
cagaaaacct caagtgaatt ttgtatgtaa atgaaatctt gaatgtaagt tctgtgattc 540
tttaagcaaa caattagctg aaaacttggg attgtttagg tttatgtagt aagtgacttg 600
gcacccatca gaaaataaag ggcattaaat tg 632
```

<210> 326

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig47563

<400> 326

```
agcaacaat tagctgaaaa cttgggtattg ttgtagttaa tgtagtaagt gacttggcac 60
```

<210> 327
 <211> 540
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig48913

<400> 327
 accagagggt gtccttttc cacagtaatg ggatoggctg gtgtgccttc agggaggaag 60
 agggaggtgg tcaagcttga aaaactggct ttaggatggg tctgactttg ttctccctcc 120
 ccaagtgttc tcaacctcca ttctgcagtg ttcagagttt tagggaaagg gtttggtgtg 180
 cccagcatcc aggtgttggt tggcttagcg catgtgaagt gaaaaccttc tggggttggt 240
 tggaagcagc tttctggttc ttgtgattgt atcctgaggt cccagaacct tattctccca 300
 cgaggatcct cagtgacct ggtggccaca cgcctggcca gcctgctggc tcctgggtga 360
 gctgaagaac cttgcctgtg gcacttttcg agggtagagt ggaaccgaga gaacatggtc 420
 cccgtgctgg gactcatgcg ggtcatttcc tgccggcctg gtttcgcctg gtcgtgtctt 480
 tatgagcacc atgtaagcct ccttgatttg agataattgg gcattaaaca ttaactgca 540

<210> 328
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig48913

<400> 328
 tatgagcacc atgtaagcct ccttgatttg agataattgg gcattaaaca ttaactgca 60

<210> 329
 <211> 534
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig49169

<400> 329
 cctaattgtta acatttttaa aaatacatat ttgggaactct tattatcaag gttctaccta 60
 tgtaatttta caattcatgt ttcaagacat ttgccaaatg tattaccgat gcctctgaaa 120
 aggggggtcac tgggtctcat agactgatat gaagtcgaca tatttatagt gcttagagac 180
 caaactaatg gaaggcagac tatttacagc ttagtatatg tgtacttaag tctatgtgaa 240
 cagagaaatg cctcccgtag tgtttgaaag cgtaaagctg ataatgtaat taacaactgc 300
 tgagagatca aagattcaac ttgccatata cctcaaattc ggagaaacag ttaatttggg 360
 caaatctaca gttctgtttt tgctactcta ttgtcattcc tgtttaatac tcactgtact 420
 tgtatttgag acaaataggc gatactgaat tttatactgt tttctacttt tccattaaaa 480
 cattggcacc tcaatgataa agaaatttaa ggtataaaat taaatgtaaa aatt 534

<210> 330
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig49169

<400> 330
 catacacctc aaattcggag aaacagttaa tttgggcaaa tctacagttc tgtttttgct 60

<210> 331

<211> 602
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig49388

<400> 331
 tgtcagtgga ggggtctctg cagccaactg agactatctt gctgtgccct gaggcttctc 60
 agggtttaga agaacagcat tcaaaattcc ccgtcctgtc agtggttgcc ttgcacctc 120
 ctcccctaaa gcagcgcggg gggcaaataa gacccaccc ctccctgcag cttcacagg 180
 acgcttcctt ccctcccgc aaccaccca ggctcccctg ggaggctgca gttgtggtac 240
 acgtccccgg tgcgtgggtg gccgtgactc gggggcggg cgatcgggtc tcagcccctg 300
 ccttccccag tctctgggtc acccgaattt tcccaccct gcttctcccc gaggagggtg 360
 agctottgag caagtggga cttgggccgg gccctggaag aatgattggc tgggaggccg 420
 cgggaggag gccaggagg ccggaccagt tgggaggagt gagcaggccc cggggagg 480
 ggatgagcgc agtttgctcg ctttctccc ctgccggccc cctccgccc cacacacact 540
 cgggacgtct tcattgaaga ttcacttaca aaggaatgtt tcactaaata aaagaaaacc 600
 ag 602

<210> 332
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig49388

<400> 332
 cgggacgtct tcattgaaga ttcacttaca aaggaatgtt tcactaaata aaagaaaacc 60

<210> 333
 <211> 562
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig50728

<400> 333
 gcgaatttgg gccccttgat cctctgatgg gagctgaaag gatgagaggt gggcatctag 60
 atttaggagg gctgttcagg ctttgcaggt cccttacctg aacacataga aacctggag 120
 ctgtgactgt gtccatgtgt gtgtgtttgt ctgtgtgtgt tgcgggggat gggcacctgc 180
 atgaatgtgg tagagaaaat ggctctgctc agagggaaga tacgcatagc aaggcaggga 240
 ccagaggaat cacaggcgcc tggagagcag ccgggcaccg cctccaggga cctgccggct 300
 tccctcagtc ctccaggggc ccagcactct tcccttaggc cctgtgagcg tcccttgtca 360
 ggatacattc tctcattttg ctgaagctga tttgattggg tgtctgttcc tcgcagccaa 420
 aagagctctg aatgaggaaa gtgcttctgt gctaactccc cgcgtctcct gaatttcagt 480
 cattcatgta cccgcctcga aatttttgca atatctgtgt accaactgtc catttactta 540
 ataaagaagt tttcttttaa tt 562

<210> 334
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig50728

<400> 334
 tttgattggg tgtctgttcc tcgcagccaa aagagctctg aatgaggaaa gtgcttctgt 60

<210> 335

<211> 400

<212> DNA

<213> Homo sapiens

<300>

<308> AI497657

<400> 335

```
tttttttttt tgcacttatg gtattttattg ttggaagatt gagtacctta atgcacacca 60
atgctcagat gacttggggg cacatagggg actgctgtca ccatgcctca ctctgcagg 120
gaaggggctg ccctactaaa accccagcgg gccagtgct gtgtccagaa caggtcctta 180
tattactgca gccacaatg gaactactga gtaggagcca aaagaggagg gagcaggaag 240
aggtggcatt tggagagggg agaccgcacc cacaggtctg ccacagcgg tcaacggat 300
ggggtacttt tacagtcaag ttgacttcgg tgtccgcca ccatctacct ttgtaggacc 360
actgaaacaa gggacatcca ccacggcca cagccggggc 400
```

<210> 336

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> AI497657

<400> 336

```
gagcattggg gtgcattaag gtactcaatc ttccaacaat aaataccata agtgcaaaaa 60
```

<210> 337

<211> 475

<212> DNA

<213> Homo sapiens

<300>

<308> Contig50950

<400> 337

```
ctggaagagg ctcccaaccc agagtgtccc tgtgggaggc aggcagaagg tgacaattga 60
cacgatttcc tgcacgcgtc ctctctacc ttggaagcag ttagaatcta ccaggcacag 120
atgaggccgc ccttgctga cggagcttga tgagcagccc ttggtctccg gttccaggac 180
tgagagccca gctgcctctg cccacccttc cccaggcctc tgccagcctc tggctgcacg 240
gtcaggccct gcccctatggc aggcctgccg gagcttggct ggggaccct cccgcctctg 300
gtcctctgat gggctggatg taacttgtgt cttctagccc cttaaggagc ccaggtgttt 360
taagggaatga attggtcact gcattcttga tcgattatgg ttctgagaaa agcaaatatc 420
acttttggct gcattaaaag aagcatcata tataaaataa agaagatgaa ggtct 475
```

<210> 338

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig50950

<400> 338

```
gtcactgcat cttgtatcga ttatggttct gagaaaagca aatatcactt ttggctgcat 60
```

<210> 339

<211> 860

<212> DNA

<213> Homo sapiens

<300>

<308> Contig51660

<400> 339

```

ggatggcaac cttcagctag actgcctggc tcaaggggtgg aagcaatacc aacagagagc 60
atttggctgg ttccgggtgtt cctcctgcca gcgaagttgg gcttccgcca agtgcagatt 120
ctgtgccaca cgtactggga gcactggaca tcccagggtc aggtgcgtat gaggtctttt 180
ggccaaaggt gccagaagtg ctccctgggtcc caatatgaga tgcctgagtt ctccctcggtat 240
agcaccatga ggattctgag caacctgggtg cagcatatac tgaagaaata ctatggaaat 300
ggcatgagga agtctccaga aatgccagta atcctggaag tgtccctgga aggatcccat 360
gacacagcca attgtgaggc atgcactttg ggcatatgtg gacaggggctt aaaaagctac 420
atgacaaagc cgtccaaatc cctactcccc cacctaaaga ctgggaattc ctcacctgga 480
attggtgctg tgtacctcgc aaaccaagcc aagaaccagt cagatgaggc aaaagaggct 540
aaggggagtg ggtatgagaa attagggccc agtcgagacc cagatccact gaacatctgt 600
gtctttatit tgcctgttgt atttattgta gtcaaagtct ttacatcaga atgatgaaaa 660
taggcttgcc actttctctt attttaattc catggtagtc aatgaactgg ctgccacttt 720
aatataactg aaaattcatt ttgagaccaa gcaggatcaa gttttagtaa taaacactgg 780
tttcttagcc atcctctgaa aacagtatga aacatgacca agtacataat ggatttagta 840
ataaatattg tcgaattgct 860

```

<210> 340

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig51660

<400> 340

```

gctgcttgta tttattgtag tcaaagtctt tacatcagaa tgatgaaaat aggcttgcca 60

```

<210> 341

<211> 608

<212> DNA

<213> Homo sapiens

<300>

<308> Contig52490

<400> 341

```

atcgtggcta gcggacagac acgagcctct tgggaatacc ttgtccatca cgtcatggcc 60
atgggtgcct tcttctccgg catcttttgg agcagctttg tcggtggggg tgtcttaaca 120
ctactggtgg aagtcagcaa catcttcctc accattcgca tgatgatgaa aatcagtaat 180
gccaggatc atctcctcta ccgggttaac aagtatgtga acctgggtcat gtactttctc 240
ttccgcctgg cccctcaggc ctacctcacc catttcttct tgcgttatgt gaaccagagg 300
accctgggca ccttctgctc gggatatctg ctcatgctgg acgtgatgat cataatctac 360
ttttccgcc tctcctgctc tgacttctgc cctgagcatg tccccaagaa gcaacacaaa 420
gacaagttct tgactgagaa ctgagtgagg ggcacagagc ctgggacaac aaaaacggac 480
aaggccagaa acagcttcat atggacactg ggacttagcc ccaagcctgg gtgtcctctg 540
aggccagcct ctccaccttc tgagcctgcg cccacactat tgaaaacact aatgaaaagta 600
ctcctctg 608

```

<210> 342

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig52490

<400> 342

ccaggatcat ctcctctacc gggtaacaa gtatgtgaac ctggtcatgt actttctctt 60

<210> 343

<211> 1282

<212> DNA

<213> Homo sapiens

<300>

<308> Contig53598

<400> 343

```

catgccagca cctttgaacc ggtctcttag aagaagacac acatcctggg tgtacagtgg 60
tgaaatgggg agtgggtgcc cattctgaaa aacgaggcat tcctgctcat tccctctgct 120
tagctggtgg gcaggggaga gagggaaatg ccaaaaactt ggagtgaagg atgatgctat 180
tttttatttt taaatatatc ttcaggttat ttcttactg ttgcttcaga tctaatagtaa 240
aaggcagatg tccccctcctc tccacccccg acgtgaccc cggcctcagt cacggctctt 300
tgcagatgca cagttctgtg ttctggcctg tggcagggcc gggaagggcc gctggcttcc 360
gaacagacgt ggttgctctc cacgaggcgc atggggagcc cgcgggccct aagctttgtc 420
gcagatgtca tcattggcag aattacttgt cttgaaaaat aagtagcatt gctgaaacac 480
acaaccgaat tctctacgat ggccatttgc tcattgtctt tcctctgtgt gtagtgagtg 540
accctggcag tgtttgccctg ctccagagtgg ccctcagaa caacagggct ggccttgga 600
aaaccccaaa acaggactgt ggtgacaact ctggtcagggt gtgatttgac atgagggccg 660
gaggcggttg ctgacggcag gactggagag gctgcgtgcc cggcactggc agcgaggctc 720
gtgtgtcccc caggcagatc tgggcacttt cccaaccag gtttatgctg ctccaggga 780
gcctcggtgc cagagtgggtg ggcagatctg accatcccca cagaccagaa acaaggaatt 840
tctgggatta cccagtcccc cttcaaccca gttgatgtaa ccacctcatt tttacaaat 900
acagaatcta ttctactcag gctatgggcc tgcctcctac tcagttattg cgagtgttgc 960
tgtccgcagt ctccggggccc cacgtggctc ctgtgctcta gatcatggtg actccccgc 1020
cctgtggttg gaatcgatgc cacggattgc aggccaaatt tcagatcgtg tttccaaaca 1080
cccttgctgt gccctttaat gggattgaaa gcacttttac cacatggaga aatatatttt 1140
taatttgtga tgcttttcta caaggtccac tatttctgag tttaatgtgt ttccaacact 1200
taaggagact ctaatgaaag ctgatgaatt ttcttttctg tccaaacaag taaaataaaa 1260
ataaaagtct atttagatgt tg 1282

```

<210> 344

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig53598

<400> 344

ccactatttc tgagtttaat gtgtttccaa cacttaagga gactctaataa aaagctgatg 60

<210> 345

<211> 601

<212> DNA

<213> Homo sapiens

<300>

<308> Contig53641

<400> 345

```

tggaggctgt ggatgatgct ttcaagacaa tggatgtgga tatggccgag gaacatgcc 60
gggccagat gagggcccag atgaatatcg gggatgaagc gctgattgga cgggtggagct 120
gggatgacat acaagtcgag ctctgacct gggatgagga cggagatttt gccgatgcct 180
gggccaggat cccctttgct ttctgggcca gataccatca gtacattctg aatagcaacc 240

```



```

gtgccaacag gagggccacg tggagagctg gcgtcagcag tggcaccaat ggagggggcca 300
gcaccagcgt cctagatggc cccagcacca gctccaccat ccggaccaga aatgctgcca 360
gagctggcgc cagcttcttc tcctggatcc agcaccgttg acgaactgca gcgatcttac 420
tggccaagcc agagcgcctc ctctcagatt ccttctcgac acagcacctt aggcggcttc 480
ttcctgtcag tcggagggtg catgcaagat gaagctctct ttgctcttcc tgctttcatt 540
ttgtgctttt ccttgtgttt tcatgttttg ggtatcagtg ttacattaaa gttgcaaaat 600
t 601

```

<210> 346
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig53641

<400> 346
 ctttcatttt gtgcttttcc ttgtgttttc atgttttggg tatcagtgtt acattaaagt 60

<210> 347
 <211> 751
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig54242

<400> 347
 aattactcaa agaaggagcc atttcagtta actcaagtga atgaaagact tttggaatct 60
 gcagtgggtc cttccctgtt gaccatttgg taacttgtaa tctgaccaa aactcttgag 120
 ctgcaacagg ccttgccaga gggctcagga tgggaaagga agaaggggat aggaaaagaa 180
 gaggtaattt tacatttccc ctttaaagta aatttttagcc aactcatcat tctgaaatgt 240
 ccctataaaag aatgagtcga actagaccag aagccagcct actccttott acatagcttc 300
 tccaacaggg gtagcaatga cctgtccact tcaaacacag ataaggcctg ccacccctcat 360
 tgggttaaag cacacgtgag acttttcagt ggctctgctg agaaggaagg cagcccagga 420
 gtcagggtat caggcattgc attgtcagtg tctgctctca gagtttacac attcaattgc 480
 ttccaagggt gaatctcctg ctctgtgaat gctatcaga cccaaaggcc aaccttgggc 540
 tgggtctatg tacgttcttc cgaagcactg atgatcaaaa ttgaagacac attcagaggt 600
 ttgattgggt gagattaact ggtgtgggtg ttgggtgatg tatgttttat ttttatgtct 660
 ttgtatgtag ttctacataa tgcaaattgt gctttctgat ggacaagacc tcataactgt 720
 gattaatatc aataaaaagg ggatgttgtg g 751

<210> 348
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig54242

<400> 348
 gtaaatTTTA gccaactcat cattctgaaa tgtccctata aagaatgagt cgaactagac 60

<210> 349
 <211> 637
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig54661

<400> 349
 ggcagtgatg tctatgttga gattaactta tgtattgagg aaaatttgaa gtttattttt 60
 tcgatgaata aggctgtcaa atgatttagt atagattaat gacatctttt ttagaaatat 120
 taaagtgagt attcctcatt atgtcatcat ttctgataat tagagtgcta atttgaatgt 180
 tagataatgt ttccacatct atacctattt ctttctaggg cacttctgac cctggggcctt 240
 ggggatggcc tttaggccac agtagtgtct gtgttaagtt cactaaatgt gtatttaagt 300
 agaaacattc ctatgtaaaa atgtgtgtat gtgaacgtat gcatacattt ttattgtgca 360
 cctgtacatt gtgaagaagt agtttgaaa tttgtaaagc acaaaccata aaagagtgtg 420
 gagttattaa atgatgtagc acaaagttaa tgttttagctt ataaaaggct ctttctattt 480
 tctatggcaa agactttgac acttgaaaaa taaaaccaat atttgattta tttttgtaag 540
 tatttaggat attattttta ataaatgatt gtccattatc aatataatag ttgtgaaatg 600
 atttaagtaa ataaacttta tgcttctgtg tctgttg 637

<210> 350

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig54661

<400> 350
 ctgtacattg tgaagaagta gtttggaat ttgtaaagca caaaccataa aagagtgtgg 60

<210> 351

<211> 924

<212> DNA

<213> Homo sapiens

<300>

<308> Contig55188

<400> 351
 gcgacaagta ccgcaagcgg gcactcatcc tgggtgtcact gctggccttt gccggcctct 60
 tcgccgccct cgtgctgtgg ctgtacatct accccattaa ctggccctgg atcgagcacc 120
 tcacctgctt ccccttcacc agccgcttct gcgagaagta tgagctggac cagggtgctgc 180
 actgaccgct gggccacacg gctgcccctc agccctgctg gaacagggtc tgcctgcgag 240
 ggctgccctc tgcagagcgc tctctgtgtg ccagagagcc agagacccaa gacagggccc 300
 gggctctgga cctgggtgcc cccctgccag gcgaggctga ctccgcgtga gatggttgg 360
 taaggcgggg tttttctggg gcgtgaggcc tgtgagatcc tgaccaagc tcaggcacac 420
 ccaaggcacc tgcctctctg agtcttgggt ctccagttcct aatatccgc tccttgcctga 480
 gaccatctcc tggggcaggg tccttttctt cccaggctcct cagcgtgccc tctgctgggtg 540
 ccttctcccc cactactact ggagcgtgcc cttgctgggg acgtggctgt gccctcagtt 600
 gccccagggt ctgggtgccc accatgcccc ttctcttttc tcctcctacc tctgccctgt 660
 gagcccatcc ataaggctct cagatgggac attgtgggaa aggctttggc catggtctgg 720
 gggcagagaa caagggggga gacacaagta gacctcaggt agaacgacac tgggcggagc 780
 caccagggt cctgtcccca gggagtgtct gaggcgcac aggcccggtt tttaccagtt 840
 tatatcacgg tcttcatttt taaaagtaac gctaaacttg tacggacgat gtctcatgga 900
 ttaaataata ttctttatgg cagt 924

<210> 352

<211> 60

<212> DNA

<213> Homo sapiens

<300>

<308> Contig55188

<400> 352
 agtaacgcta actttgtacg gacgatgtct catggattaa ataattttct ttatggcagt 60

<210> 353
 <211> 699
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig55353

<400> 353
 tgattatgcc aagagctcta aacagaagtt tgagaaggta aaaattaagt tgtagtatct 60
 gagttgtttt ttttttcttc ctttgggtgt tatgaaggta ttcataagaa ctttaatttc 120
 aggggaaaaa atgcctgatt tgctattttt gacatttctt cgtctcttaa gaagtcagtt 180
 aaatatgttt tcatagttta tattcctgtt tcatagatta ctgtgaaaca tgtattttaa 240
 cctatgaatt ataaaatagt atttagattc tagcgtgagt taaatagatt agtcatatat 300
 ctttttagatt tgtggatttg acatgtaaat tatgtgttgt gtataagtaa gttagttact 360

 aaacatatgg catggttatt gataaacttg ttgctatttt tttccaaatg ctatcagttg 420
 ttgtggactt ttaaaaaatta gtttgaattt tggaatgttc tgtgataaaa tataatttca 480
 actattttgt acattttaa atgccatgtt gtatatgtct gtatttaaaa atgttgtaaa 540
 tatctgcatt ttaagaattt tgaaagattt tcctcaaaaa tgacagaact ctccatactt 600
 aattgtgaca cattataaga tatctgattt taagcttttg gattttgttc taaaaattaa 660
 gtttaaacat gctgaaaatt ccataaaaaat aaaattttg 699

<210> 354
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig55353

<400> 354
 taaaatagta ttttagattct agcgtgagtt aaatagatta gtcatatatc ttttagattt 60

<210> 355
 <211> 809
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig56503

<400> 355
 gcatgtgaga tgagtgactg ccggtgaatg tgtccacagt tgagagggtg gagcaggatg 60
 aggggaatcct gtcaccatca ataatcactt gtggagcgcc actctgcca agacgccacc 120
 tggggcgaca gcatggagct ctccatggcc aggcgtgctg tgtgcatgtt cctgtgtctg 180
 tgcccccttg ccgcctcctt gcaaacctca cagggtcccc acacaacagt gccctccaga 240
 agcagccctt cggaggcaga ggaaggaaaa tggggatggc tggggctctc tccatcctcc 300
 ttttctcctt gccttcgcat ggctggcctt cccctccaaa acctccattc cctgtgtgcc 360
 agcccccttg ccatagcctg attttgggga ggagggaagg gcgattttgag ggagaagggg 420
 agaaagctta tggctgggtc tggtttcttc ctttcccaga gggctcttact gttccagggg 480
 ggccccaggg caggcagggg ccacactatg cctgcgcctt ggtaaagggt acccctgcc 540
 tttaccagca gccctggcat gttcctgccc cacaggaata gaatggaggg agctccagaa 600
 actttccatc ccaaaggcag tctccgtggt tgaagcagac tggatttttg ctctgccctt 660
 gacccttgt ccctctttga gggaggggag ctatgctagg actccaacct cagggactcg 720

 ggtggcctgc gctagcttct tttgatactg aaaactttta aggtgggagg gtggcaaggg 780
 atgtgcttaa taaatcaatt ccaagcctc 809

<210> 356

<211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig56503

<400> 356
 gaaaactttt aagggtgggag ggtggcaagg gatgtgctta ataaatcaat tccaagcctc 60

<210> 357
 <211> 976
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig56678

<400> 357
 gaaggatata ctttggtata acttattatt ttgttctctg taaatacaag atgtttatag 60
 gaaatatgta ttctgaactc tatctgcaga atgagtcact acaccaaaat agttctatta 120
 tttagaatgt gtttaatttta aagggaacctg atagggtattt atttacaat gcgatccaca 180
 tttgtgtgaa agcatgtgat cataactaacc cagcctcctg gaatgtcgt gtacgatgat 240
 tgatgtcttt ttctcagtc atagttacaa ttgtttagta tgctaatacag tccagttccc 300
 tgagggttaa gatcaaata aaattactct gcttttcgac tcattcaggt agcattgtac 360
 ctgaacctga ttgctacttt ttcactctaa atattatatt tcctcatcta atctgccttc 420
 ccctcatcca cagacatttg gagaaggaaa tgggaggggtg tctgttatcc ctttctcttt 480
 gctttgtccc cgttggttaga ctggcagcgt cagttgctcg gtgggcttgg ttagagccgt 540
 ggggtaggca ggtggctggc ggggacagg agaggctgag agggaagtgg tggcatttac 600
 tgctctgaca cttccactgt ccctgctggg gatgctgggg ccaaggcctg tggggcctgt 660
 gaactgcaca gccaggagca aggaaccac taaatactcc gtcaacctca tgtcccctct 720
 acagtgttaa attattacat aagcagggtga aaggtagaag gcgaattatg tgagtaaata 780
 tgggtctgttt tctcttcagc aaaaatgact atttttgtgt gtgactaatt tatttttatt 840
 attgtaaaga tacaataaac cggttgaaat atctgctttg ttgacaagcg tgtgctttct 900
 ctggccttat tcgcgttctg ttctcctgca aatagcggcc tctaaaaaga agagtcagac 960
 aataaactgg ttgaaa 976

<210> 358
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig56678

<400> 358
 tattacataa gcagggtgaaa ggtagaaggc gaattatgtg agtaaataatg gtctgttttc 60

<210> 359
 <211> 1118
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig57584

<400> 359
 agctgttgtg catccagagg tggaaattggg gcccggcatt ccctcctcgt cccgggctgg 60
 cccttgcccc caccctgcaa ctctgggttg agatgggctc agccaagagc gtcccagtc 120
 caccagcgcg gcctccggcg acaacaagca tctggctcga gtggcgagacc cccgttcacc 180
 tagtgctggc atcctgcgca ctcccatcca ggtggagagc tctccacagc caggcctacc 240

```

agcaggggag caactggagg gtcttaaaca tgcccaggac tcagatcccc gctctcctac 300
tcttggtatt gcacggacac ctatgaagac cagcagtgga gaccccccaa gccactgggt 360

gaaacagctg agtgaagtat ttgaaactga agactctaaa tcaaattcttc ccccagagcc 420
tgttctgccc ccagaggcac ctttatcttc tgaattggac ttgcctctgg gtaccagtt 480
atctgttgag gaacagatgc caccttgga ccagactgag ttccctcca aacaggtgtt 540
ttccaaggag gaagcaagac agcccacaga aacccctgtg gccagccaga gctccgacaa 600
gccctcaagg gaccctgaga ctcccagatc ttccaggttct atgcgcaata gatggaaacc 660
aaacacgagc aaggtactag ggagatcccc cctcaccatc ctgcaggatg acaactcccc 720
tggcaccctg acactacgac agggtaagcg gccttcaccc ctaagtgaat atgttagtga 780
actaaaggaa ggagccattc ttggaactgg acgacttctg aaaactggag gacgagcatg 840
ggagcaaggc caggaccatg acaaggaaaa tcagcacttt cccttggtgg agagctaggc 900
cctgcatggc cccagcaatg cagtcaccca gggcctgggt atatctgtgt cctctcacc 960
cttctttccc agggatactg aggaatggct tgttttctta gactcctcct cagctaccaa 1020
actgggactc acagctttat tgggctttct ttgtgtcttg tgtgtttctt ttatattaaa 1080
ggaagtaatt ttaaatgtta ctttaaaaag gtatatgt 1118

```

<210> 360
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig57584

```

<400> 360
aggaatggct tgttttctta gactcctcct cagctaccaa actgggactc acagctttat 60

```

<210> 361
 <211> 859
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig63649

```

<400> 361
gtcgcagggt accagtgtgc ggagttcctg ttgccaaagt gaaggtggcc ctgggcaggc 60
acaggtgtgg tcatatcttc agccaacagg accatcctcc ggagggccac ctctggggac 120
ttcctacggg aagagagtga cagatttggt gcttctgtgt gtttctgccg cttcagtgagg 180
gccgtgctgg gagacagcgg gtggatcctc cagcagcctg tctgctgagc ctgccttctc 240
aagtctactg ttaaaatcag gaccgggtcg tgtccgagcc tacaggccct gtctccgctc 300
ccaggcctg caggagttga gggctgcacc tgctcgctgg agaggagag gcagatttag 360
tggacgcctg gcatggactc ggactggcct ttggaagctc cctgccctga cgggttgcc 420
gtcaccactg cgaagtgagg cttggaggac ctgcacctga gaaaggctgt gtgtggtctt 480
gggtccacac ctgccagagc taacttactg ccagacggcg acttactgtg ggccaccctc 540
agtgaaccgg ggtgtcctca gctggcccta cagagcactt ctgtgctggg gatgagtagg 600
aactctgggc gaggagggtc ccagcgccgc ccctcgatac agocctgtct tgccctctgc 660
cgtacttat accaggtggg atccctgccc tgcattgcct ggggattggc tgggcttggg 720
cacgccctgc tgtggaactg gatgttttca gggagcccag cctttcctca tgtcaacaca 780
gttacaata tagttttcaa agtacagttt aaaactcaaa agtaaaactt tcagcaactc 840
aaaaaaaaa aaaaaaaaaa 859

```

<210> 362
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig63649

<400> 362
cagcctttcc tcatgtcaac acagttcaca atatatgtttt caaagtacag tttaaaactc 60

<210> 363
<211> 1170
<212> DNA
<213> Homo sapiens

<300>
<308> Contig63525

<400> 363
gccatggctc cctggggcgga gcgagcactc gcggtgaac ccgctgcgcg cgggtgtggct 60
cacgctgacc gccgccttcc tgcctgacct actgctgcag ctccctgccg ccggcctgct 120
cccgggctgc gcgatcttcc aggacctgat ccgctatggg aaaaccaagt gtggggagcc 180
gtcgcgcccc gccgcctgcc gagcctttga tgtccccaag agatattttt ccactttta 240
tatcatctca gtgctgtgga atggcttccct gctttggtgc cttactcaat ctctgttccct 300
gggagcacct tttccaagct ggcttcatgg tttgctcaga attctcgggg cggcacagtt 360
ccagggaggg gagctggcac tgcctgcatt cttagtgcga gtatttctgt ggctgcacag 420
cttacgaaga ctcttcagat gcctctacgt cagtgtcttc tccaatgtca tgattcacgt 480
cgtgcagtac tgttttggac ttgtctatta tgccttgtt ggcctaactg tgotgagcca 540
agtgccaatg gatggcagga atgctacata acagggaaaa atctattgat gcaagcacgg 600
tggttcata ttcttgggat gatgatgttc atctggtcac ctgcccatca gtataagtgc 660
catgttattc tcggcaatct caggaaaaat aaagcaggag tggtcattca ctgtaaccac 720
aggatcccat ttggagactg gtttgaatat gtttcttccc ctaactactt agcagagctg 780
atgatctacg tttccatggc cgtcaccttt gggttccaca acttaacttg gtggctagtg 840
gtgacaaatg tcttctttaa tcaggccctg tctgccttcc tcagccacca attctacaaa 900
agcaaatattg tctcttacct gaagcatagg aaagctttcc taccattttt gttttaagtt 960
aacctcagtc atgaagaatg caaaccaggt gatggtttca atgcctaagg acagtgaagt 1020
ctggagccca aagtacagtt tcagcaaagc tggttgaaac tctccattcc atttctatac 1080
cccacaagtt ttactgaat gagcatgcag tgccactcaa gaaaatgaat ctccaaagta 1140
tcttcaaaga attaattact aatggcagat 1170

<210> 364
<211> 60
<212> DNA
<213> Homo sapiens

<300>
<308> Contig63525

<400> 364
ctcttaccgg aagcatagga aagcttttcc accatttttg ttttaagtta acctcagtc 60

<210> 365
<211> 632
<212> DNA
<213> Homo sapiens

<300>
<308> Contig64688

<400> 365
aagaatgcta agatgatttc agatatcgaa aagaaaaggc agcgtatgat tgaagtccag 60
gatgaactgc ttogggttaga gccacagctg aaacaactac aaacaaaata tgatgaactt 120
aaagagagaa agtcttccct taggaatgca gcataattct tatctaattt aaaacagctt 180
tatcaagatt attcagatgt tcaagctcaa gaaccaaacy taaaggaaac gtatgattca 240
tccagccttc cagctctgtt atttaaagca agaacacttc tgggagccga aagccatctg 300
cgaaatatca accatcagtt agagaagctc cttgaccagg gatgagaaga gcagtctact 360
aaaatgtgcc tataggaaga ctagtctcat gctgttaoct tctgaaactg tacctttata 420
aatcaattgt tttgcaaaga agttatggcc tacttagaat ctaaaatttg ttattcaaat 480

taaatggctg tgaacaatgt taaatagcat cagtttgtcc aatagtttta aaggccataa 540
 tcatcttttc tggttaatat cttgagtaat tttaaaatgt tgacacotta atcgggccca 600
 ggtatgagcc ataataaact tgtaaaatta ag 632

<210> 366
 <211> 60
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig64688

<400> 366
 ggctgtgaac aatgttaaat agcatcagtt tgtccaatag ttttaaaggc cataatcatc 60

<210> 367
 <211> 2596
 <212> DNA
 <213> Homo Sapiens

<300>
 <308> AL080059

<400> 367
 gatgtcacct gctttatttc tggctttggc ctgtggtctg tgatacccat cctgcttgat 60
 gttctgcaga atggcacttg actgctgggc atgcatgaag ttaagggcaa gaaacagtat 120
 gccatgtggt ctgtaccatc atgtgtctct tcttgcttct gggcccttct actggtgaac 180
 tttcatcaag atctgcgcca tgcctgttca ctatcaagcc attaatgttt gtctgggttg 240
 ctgtcagccc cagttggctt cctggccaac aaggacctca agaactgcct gtggaccgag 300
 gcccttacc agtgcatgag acacacacct accctcccca gctttccagg aaccctactg 360
 gctgccagac tgatggggcg gctggtatgt gtggacatgt gttcactgtc attatgctgt 420
 ggctccaggt gaggtgagg actgggccta tatagaatcc agataccatt gtcaacttcc 480
 cttattcccg tctaagatgt gagcagagt ccatagtagg ggttctggga agaggtat 540
 ctgatttctg ggcctctgct tgcttgactt caggtcactt atacttctta ttttgcttgc 600
 ctgccttcat ccctcatttc ctccctctca ttcttcttct ctccctccct ttcttggtag 660
 cctcctttcc tcccttctg ccttccctt ccttcttctc ttattctttt ttattttgtt 720
 taaatagtac cacagagaaa acaactgaaa aaccacattt ttctacatac agctggggag 780
 gtagctgaga acttggcact gcgcacacat actaggttga aagagagttg aggaaaccag 840
 aaggccaagt ggatctgctg gcaaaccctg aacctgtctc ctgcgcttgc tctacagttc 900
 tgaagttgaa aatcgttttc atgcctagca tctgcttgag ttataaacc caaggcagcc 960
 atgtcataga ctagtgttta ctcttgtttt gactttgttt taatgcttcc taagacccaa 1020
 gtgcctcctg ctgtttcctc ctttgggta gctctggcc atctggacct caatccccag 1080
 ctttccact ttcagcagtc ctttgcctc tttgcttcta cctcaaatag cccaggagt 1140
 gggctttagt ctccaatatg gagcatctca agcttctcct gggggatggg gattgggatg 1200
 ggcggaatct gttttggatc tccgggttat ttccagtggg tgtaaaagca gagctgggac 1260
 tttccctctc ttatccctga ggggtggtaa gaaggactgt atctacacct gttcttccct 1320
 acctctctt ttgttaggga ggcctcattc taagttcctc aagagagtcc ttggcttaaa 1380
 gctgtagcaa ggtgtgcta ggtgggggat ttggagcaaa accgtcagat aggcattgata 1440
 ctggtatgga gtgggcctgc aaaatcagac agaaatggct tgagaagccg cagggggagc 1500
 atgctgtct ctactgata gagtatggga gggacctccc tagcttgga aatgagaatt 1560
 gaaggggtta tgaacaaata ggatgcctag ttgaggatgt tcccaaagt ttgtccaatc 1620
 ttatcattag tagattttat aagccacaga gacaaaccag aaacggaata atgttacttt 1680
 ggatgcttta tttttttgtt ctagggtgtg ctttgtacat gcagaagaat gctatatgct 1740
 gcacattttg cctttaaagt cttacgactt tccccatttt agtctaattg gaagatacag 1800
 atgtgcaagt ctgctttttt gttttttgtt attatttttt tttttgctct gtgttatgga 1860
 cattttcaga catgcacaga agtggagagg atggctcctg gaccccatgt gtccatcacc 1920
 tagctgcac acttatcagc tatggtcaac ctggtttcat ctgtatctct ctcttttccac 1980
 ctgtattgtt tattgaaat ccaagacact atgccaatgc aaccgtgact actttgggag 2040
 attggtagtc tcttttgatg gtgatatgta tgggggtgcac tatcataatc acatcaggtc 2100
 tgctttttgc ttttaatgtt aactaatgaa gttccagaga tgggccttag aaatgtgttt 2160

```

taagaattaa caaggagtct caaaaagaaa tgagagggat gcttcctttc ccttgcatct 2220
acaaaacaag agagagactg ttctgttgta aaactctttc aaaaattctg atatggtaag 2280
gtacttgaga cccttcacca gaatgtcaat ctttttttct gtgtaacatg gaaacttggtg 2340
tgaccattag cattgttatc agcttgtagt ggtctcataa ctctggtttt ggaagaataa 2400
tttgaaattt gttgctgtgt tctgtgaaaa taacctcccc aaaataatta gtaactgggtt 2460
gttctacttg gtaatttgac accctgttaa taacgcaatt atttctgtgt tcttaaacag 2520
tataaatagt tgtaagtttg catgcatgat ggaaaaataa aaacctgtat ctctgtcaaa 2580
aaaaaaaaaa aaaaaa

```

<210> 368
 <211> 443
 <212> DNA
 <213> Homo Sapiens

<300>
 <308> AA555029

```

<400> 368
tttttttttt tttttccatt aaccaaacac aagtttatatt actaagtcaa gtcaaactga 60
ggagtatttg tttcttttgt agttggtaga caaagaatac atatacatat tctatttccc 120
attaagcatt cgatcatggt acaaaacaat ggccatagaga actatcattg aagatttacc 180
aaatctgcct gaagcaaaat gtgataaact gcagaaatgg tggagaatga cactggaagt 240
cataaagtgt attctcaaac acgggtttga actgcatagg accgcttata tgcatatttt 300
gataaatata ttgaaaattt ttttggaat ttgcaacaat ttaaaaaact tgcagatgca 360
ccatgtagct tagaaatact gaaaaactaa gaaaaagggtg tgtcatgaat tcatataata 420
tatgtagaca ctagtctatt tta

```

<210> 369
 <211> 3160
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003748

```

<400> 369
attctcgccg cgcccgggcg gacgatccag cgaacagccc cgcttctaac ccgagatgct 60
gctgcggcg cccgcgctcc gccgcgccct gctgtccgc ccctggaccg gggccggcct 120
gcggtggaag cacacctcct ccctgaagggt ggccaacgag ccctgtcttag ccttcacgca 180
gggcagccct gagcgagatg ccctgcaaaa ggccctgaag gacctgaagg gccggatgga 240
agccatccca tgcgtggttg gggatgagga ggtgtggacg tcggacgtgc agtaccaagt 300
gtcgcccttt aacctgggac ataagggtggc caagttctgt tatgcagaca agagcctgct 360
caacaaagcc attgaggctg ccctggctgc ccggaagag tgggacctga agcctattgc 420
agaccgggcc cagatcttcc tgaaggcgcc agacatgctg agtggggccg gcagggtgca 480
gatcctcgcc aagaccatgg tgggacaggg taagaccgtg atccaagcgg agattgacgc 540
tgcagcgga ctcactgact tcttccggtt caatgccaaag tatgcggtgg agctggaggg 600
gcagcagccc atcagcgtgc ccccgagcac caacagcac gtgtaccggg gtctggaggg 660
cttcgtggcg gccatctcgc cctttaactt cactgcaatc ggccggcaac tggcgggggc 720
accggccctg atgggcaacg tggctctatg gaagcccagt gacctgcca tgctggccag 780
ctatgctgtc taccgcatcc ttccggaggc tggcctgccc cccaacatca tccagtttgt 840
gccagctgat gggccctat ttggggacac tgtcaccagc tcagagcacc tctgtggcat 900
caacttcaca ggcagtgtgc ccaccttcaa aagcaggtgg ccagaaact 960
ggaccggttc cacaccttcc cagcctggc tggagagtgc ggcggaaga acttccactt 1020
cgtgcaccgc tcggccgacg tggagagcgt ggtgagcggg accctccgct cagccttcga 1080
gtacggtggc cagaagtgtt ccgcctgtc gctctctac gtgccgact cgctgtggcc 1140
gcagatcaaa gggcggtgct tggaggagca cagtccgcat aaagtggcg accctgcaga 1200
ggattttggg accttcttct ctgcagtgat tgatgccaag tcctttgccc gtatcaagaa 1260
gtggttgga cagcgcgct cctcaccatc ctggctgggg gcaagtgtga 1320
tgactccgtg ggctactttg tggagccctg catcgtggag agcaaggacc ctcaggagcc 1380
catcatgaag gaggagatct tcgggcctgt actgtctgtg tacgtctacc cggacgcaa 1440
gtacaaggag acgctgcagc tggttgacag caccaccagc tatggcctca cgggggcagt 1500

```



```

gttctcccag gataaggacg tcgtgcagga ggccacaaag gtgctgagga atgctgccgg 1560
caacttctac atcaacgaca agtccactgg ctogatagtg ggccagcagc cctttggggg 1620
ggcccagacc tctggaacca atgacaagcc agggggccca cactacatcc tgcgctggac 1680
gtcgccgagc gtcatcaagg agacacataa gcccttgggg gactggagct acgcgtacat 1740
gcagtgcagc cctctcgggc tccaccgtcc agctgtctgt ccgtccaggt ggccgacctc 1800
actgcacaga cccactcca gccctccac cccttcttca tgcacagctg cctttctata 1860
atccgggctt gactcccttc ttaccactgt attctggcct ctcccatgcc tcaggctctg 1920
gtttgagatc gtgctgggga ggaacatggc cactaccctc tatcccatcg gccatgtggg 1980
aggtatgacc ctgggtgcctg gcaggttctc cctctgccct ccactgggcc cagtggctca 2040
gggacctggg gaaaggagat ggagcagctc ttggggtcct ttggggaaaa ggaggccatt 2100
ctggggccct tggcaaacct caccactcac agaggctcct ggccttgatc cctgccctc 2160
caggtgtcca gggtaaagtg taactcagac tgacctgtgg ggcacagggg gcaccagctg 2220
gccttgccct ctctggtctg ggctgtctac ctctcctact gtatctttgc ccagaccac 2280
ctggggccag agggccctgt cccagccac acacctaga tgctggcatg ccttactcca 2340
ggtgcctgtg tttggccgag gctgtgtga ttccgggtct gcaccacatg gcggggttgg 2400
ggggccgctg gaggccacct gccaaaggct gggatgggat ggtcctgccg gtttaggccg 2460
tgattctgga aaaccttga tgggccttcg tcctatgtca gccttccctt tgatcctcag 2520

```

```

gccctacctg tagagacctc cactcctaga gccagtctca gggcttgagg tttccctgca 2580
ggagctcagc caccactgtg ccatgggtgac acaggccaag gcagacattg gccctccctt 2640
ctcccagccc ccagaggcct ggcttgggt tcgtcagcat gggccgagga cgttgccctgt 2700
agaatcctoc tctgcctggg agtggctctg tgtggaccag tccctcactg gccattctt 2760
tttttgacgc agccaatctg tgaccacgat tcctccaca gatgcctcct gcttgattc 2820
tgagtggta gagatctgta aagcatgact ttcaaggatg gttcttaggg gactgtgaaa 2880
gtgttgggtc ttctccagg atgcctgcat gggacccac ccggagctgg tgtggccatt 2940
cccaagtgc cactggcca tggatggggg tgggtgctgg tgccagctgg gctgggtgtg 3000
ggttctgtgt ccttccagga tatgtgtcat ttcccatgag gggccggggc aggtggctgg 3060
gtggggcac aggtggagt attcttagtt ctactgttc tacactgtga ggtggcaatg 3120
ggatttgctc agatgccacc caataaaatg cctgttactt 3160

```

<210> 370

<211> 1546

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003862

<400> 370

```

cacggccgga gagacgcgga ggaggagaca tgagccggcg ggcgcccaga cggagcggcc 60
gtgacgcttt cgcgtgcag ccgcgcgcc cgaccccgga gcgctgacct ctggcccccac 120
gcagctccgc gccggggcgg gagagcgcaa ctgggtctcc agacccgccc cgcgtgctgt 180
cccggactg agcggggcag ccagcctccc acggacgccc ggacggccgg ccggccagca 240
gtgagcgagc ttccccgcac cgcctgcac cctcctgcac agcggctgcc gccccgcagc 300
ccctgcgcca gcccgagggg cgcagcgctc gggaggagcc gcgcggggcg ctgatgccg 360
agggcgcgcc gcggagcgcc ccggagcagc agagtctgca gcagcagcag ccggcgagga 420
gggagcagca gcagcgggg cggcgcgccg ggcggcgccg gaggcgcccc gtcccgcccg 480
cgcggagcgg acatgtgcag gctgggctag gagccgccgc ctccctccc cccagcgatg 540
tattcagcgc cctccgcctg cacttgccctg tgtttacact tcctgctgct gtgcttccag 600
gtacaggtgc tggttgccga ggagaacgtg gacttccgca tccacgtgga gaaccagacg 660
cgggctcggg accgatgtgag ccgtaagcag ctgcggctgt accagctcta cagccggacc 720
agtgggaaac acatccaggt cctgggcccg aggatcagtg ccgcggcgga ggatggggac 780
aagtatgccc agctcctagt ggagacagac accttcggta gtcaagtcgg gatcaaggc 840
aaggagacgg aattctacct gtgcatgaac cgcaaaggca agctcgtggg gaagcccgat 900
ggcaccagca aggagtgtgt gttcatcgag aaggttctgg agaacaacta cacggccctg 960
atgtcggcta agtactccgg ctggtacgtg ggcttcacca agaagggcg gccgcggaag 1020
ggccccaga cccgggagaa ccagcaggac gtgcatttca tgaagcgcta cccaagggg 1080
cagccggagc ttcaagaagc cttaagtag acgacgggtg ccaagaggtc ccgtcggatc 1140
cggcccacac accctgccta ggccacccc ccgcggcccc tcaggtcgcc ctggccacac 1200
tcacactccc agaaaactgc atcagaggaa tatttttaca tgaaaaataa ggaagaagct 1260
ctatttttgt acattgtgtt taaaagaaga caaaaactga accaaaactc ttggggggag 1320

```

```

gggtgataag gattttattg ttgacttgaa acccccgatg acaaaagact cacgcaaagg 1380
gactgtagtc aaccacaggg tgcttgcttc tctctaggaa cagacaactc taaactcgtc 1440
cccagaggag gacttgaatg aggaaaccaa cactttgaga aaccaaagtc ctttttccca 1500
aaggttctga aaggaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 1546

```

```

<210> 371
<211> 376
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig32125_RC

```

```

<400> 371
caatttgat ttttctcaat ataatgttgc aagatgtcat tatattctca ttaccataa 60
cttcccatcc aaatcttaat ggcacatttg gtattttttc aaatggcttt gagatataa 120
tcacatacca cacaattcat gtatctaaag tatacagttc ttagttctcg tggttggtgg 180
ttacatccac cagactgagg gctccctgag ggttggtgcc tcagctttcc cttcacttac 240
tcattcacca aacattcagc gcctactgag tactgggggt atgatggcca acaggagaga 300
cagtctctgc ctactgtttg gtgggggtgg agtacttagt aaccgtcgtc attattgagt 360
gcttacctgt gctggg 376

```

```

<210> 372
<211> 1639
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> U82987

```

```

<400> 372
ccgccccctc cggcgtgttc atgcccccg ggcgccaggg agcgccatgg cccgcgcacg 60
ccaggagggc agctccccgg agcccgtaga gggcctggcc cgcgacggcc cgcgccccct 120
cccgctcggc cgcctggtgc cctcggcagt gtccctgccc ctctgcgagc ccggcctggc 180
tgccgcccc cccgccccca cctgctgcc cgctgcctac ctctgcgccc ccaccgcccc 240
accgcccgtc accgcccggc tgggggggtc ccgctggcct ggggggtccc gcagccggcc 300
ccgaggcccg cgcgccgacg gtcctcagcc ctgcctctcg ctggcggagc agcacctgga 360
gtcgcccggt cccagcgccc cgggggctct ggcggggcgt cccaccagg cggccccggg 420
agtccgccc gagagggaac agtgggccc ggagatcggg gccagctgc ggcggatggc 480
ggacgacctc aacgcacagt acgagcggcg gagacaagag gagcagcagc ggcaccgccc 540
ctcaccttgg agggctctgt acaatctcat catgggactc ctgcccttac ccagggggcca 600
cagagcccc gagatggagc ccaattaggt gcctgcaccc gcccggtgga cgtcaggggc 660
tcggggggga ggccccctcc acctcctgac accctggcca gcgcggggga ctttctctgc 720
accatgtagc atactggact cccagccctg cctgtcccgg ggcggggcgg ggcagccac 780
tccagcccc gcccagcctg ggggtgactg actgagatgc ggactcctgg gtccctggcc 840
aagaagccag gagagggacg gctgatggac tcagcatcgg aaggtggcgg tgaccgaggg 900
ggtggggact gagccgccc cctctgccc ccaccacat ctccaggaaag gctgttgtgc 960
tggtgcccgt tccagctgca ggggtgacac tggggggggg gggggctctc ctctcggtgc 1020
tccttcactc tgggcctggc ctccagcccc tgggtgcttc cccctcctc ctgggagggg 1080
gcccgtgaag agcaaatgag ccaaactgta ccaactagct cctggagcca gagagtgggg 1140
ctcgtttgcc ggttgctcca gccggcgccc cagccatctt ccctgagcca gccggcgggg 1200
ggtgggcatg cctgcctcac ctccatcagg ggtggccag gaggggccc gactgtgaat 1260
cctgtgctct gcccgtagc gccccccgcc ccatcaatcc cattgcatag gtttagagag 1320
agcacgtgtg accactggca ttcatttggg ggggtgggga ttttggtgta agccgcccc 1380
gccttagtcc ccagggccaa gcgctggggg gaagacgggg agtcaggggg ggggggaaat 1440
ctcggaagag ggaggagtct gggagtgggg agggatggcc cagcctgtaa gatactgtat 1500
atgcgtgtgt gtagataccg gaatgaattt tctgtacatg tttgggttaat tttttttgta 1560
catgattttt gtatgtttcc ttttcaataa aatcagattg gaacagtgga aaaaaaggcc 1620
ttcgtggcct cgagagatc 1639

```

```

<210> 373

```

<211> 2273

<212> DNA

<213> Homo sapiens

<300>

<308> AL137718

<400> 373

```

ggggggccagt taccatattg tccggagtggt gtccttgaca ttcaggaaag atacatcctc 60
agacctttca aaatttgtca tgccctctcta acctaaaaca gccctcatta aatgcacttt 120
aatccgagca ctgtatggct tggaaagaat tatgagttag gagaggagcc tttccttatt 180
ggccaaagcc gtggatccca gacaccccaa tatgatgaca gatgtgggta aacttctctc 240
tgcggtatgc attgtagggg aagaaagcat ccttgaagaa gttttagaag ctttaacttc 300
agctggtgaa gaaaaaaaaa ttgacagatt tttttgtatt gtggaaggcc tccggcacia 360
ttcagttcaa ctgcaagtag cttgtatgca gctcatcaat gccctgggta catctcctga 420
tgatttggat ttcaggcttc acatcagaaa tgaatttatg cgttgtggat tgaaagagat 480
attgccaaat ttaaaaagca ttaagaatga tggcctggat atccaactta aagtctttga 540
tgagcataaa gaagaagatt tgtttgagtt atcccatcgc cttgaagata ttagagctga 600
acttgatgaa gcatatgatg tttacaacat ggtgtggagc acagttaaag aaactagagc 660
agagggatat tttatttcta ttcttcagca tcttttgctg attcgaaatg attattttat 720
aaggcaacaa tacttcaaat taattgatga gtgtgtatcc cagattgtat tgcatagaga 780
tggaatggat ccagacttca catatcgaaa aagactagat ttagatttaa cccagtttgt 840
agacatttgc atagatcaag caaaactaga agagtttgaa gagaaagcat cagaacttta 900
caagaaattt gaaaaagagt ttaccgacca ccaagaaact caggctgaat tgcagaaaaa 960
agaggcaaaag ataatgagc ttcaagcaga gctacaagct ttaagtctc agtttggtgc 1020
cttgccagct gattgtaata ttcccttgcc tccctctaaa gaagggtgaa ctggccactc 1080
agcacttcct cctccgcctc cactgccttc tgggtggagg gtgccgcctc cacctcctcc 1140
cccaccacct cctccacttc caggaatgcg gatgccattc agtggctctg tgcctccacc 1200
acctccctcg ggattccttg gaggacaaaa ttctcctcct ctaccaatcc tgccatttgg 1260
gttgaaacca aagaaagaat ttaaacctga aatcagcatg agaagattga attggttaa 1320
gatcagacct catgaaatga ctgaaaactg tttctggata aaagtaaagta aaaataagta 1380
tgaaaacgtg gatttgcttt gtaaacctga gaatacattt tgttgccaac aaaaagagag 1440
aagagaagag gaagatattg aagagaagaa atcgattaag aaaaaatta aagaacttaa 1500
gttttttagat tctaaaattg ccagaaacct ttcaatcttc ctgagctctt ttcgggtgcc 1560
atatgaggaa atcagaatga tgatatttga agtagatgaa acacggttgg cagagtctat 1620
gattcagaac ttaataaagc atcttcctga tcaagagcaa ttaaatcat tgtctcagtt 1680
caagagttaa tatgtgaacc tgagcagttt gtggttgtga tgagcaatgt 1740
gaagagacta cggccacggc tcagtgcctt tctctttaag cttcagtttg aagagcagg 1800
gaacaacatc aaacctgaca tcatggctgt cagtactgcc tgcgaagaga taaagaagag 1860
caaaagcttt agcaagttgc tggaaacttg attgctaag ggaaactaca tgaatgctgg 1920
ctcccggaat gctcaaacct tcggatttaa ccttagctct ctctgtaaac taaaggacac 1980
aaaatcagca gatcagaaaa caacgctact tcatttcctg gtagaaatat gtgaagagaa 2040
gtaccctgat atactgaatt ttgtggatga tttggaacct ttagacaaaag ctagtaagg 2100
tttgtgcctt ttttaagaaac atttcatggc tctgatcttc agtgctaaaa gactgaaaat 2160
tatacctttt atttgtatgt actttccact ttctcacagt gtttccatac ccaatatctc 2220
attttaaaaa taaagagagt tgtaggatat taaaaaaaaa aaaaaaaaaa aag 2273

```

<210> 374

<211> 2782

<212> DNA

<213> Homo sapiens

<300>

<308> AB037863

<400> 374

```

gcggccgcgg cggggcggcg cggggaaccgg gccccggggg gagtgcggcg ggctgctgct 60
gctgtgctc caggtgctgc cgcccggggc tacggcaggg ccgggacgcc ggggcacgcg 120
gggcgctggc ggccggcgcg tctgctggcc cggcgcggcc ccagccttcc cccgggacgc 180
gcggctgctg ctgctcctgc cgccgctgcc gccactgtcg ccgcccgcgc cgagctccgc 240
gccgcagcc tccgcctccc ggatggacgc tctgccccgc agcgggctga acctgaagga 300

```

```

ggagccgctg ctgcccgcg gcctgggctc agtgcgctcc tggatgcagg gcgcgggcat 360
cctggacgcc agcaccgcg gcgagagtgg cgtgggtctg gcacgagcac attttgagaa 420
gcagcctccc tccaacctca ggaaatccaa cttcttccac ttcgtgctgg ccatgtacga 480
ccggcagggg cagcccgtgg aggtggagcg cacagccttc atcgacttcg tggaaaagga 540
ccgagagccc ggggcggaaa agactaacia tgggatccat taccgcctcc ggctggtgta 600
taacaatgga ctgcgagac agcaagacct ctacgtgcgt ctcatcgact ccatgtccaa 660
acaggccatc atctatgagg ggcaggacaa gaaccccgaa atgtgccgag tgctgctcac 720
ccatgagatc atgtgcagcc ggtgctgtga ccggaagagc tgtggcaacc ggaatgagac 780
gccctcagac ccctcatca ttgacagggt cttcctcaag ttcttctca aatgcaacca 840
gaactgcctg aagaatgcgg ggaatccag agacatgcgc cgcttccagg tgggtggtgc 900
cacgacggtg agcgtggacg gacacgtgct ggccgtgtcc gacaacatgt ttgtgcaca 960
caactccaag catggccgca gggcgcgccg cctggacccc tccgaagctg ccacccctg 1020
catcaaggcc atcagccccg gggagggctg gaccacgggc ggcgccaccg tcattgtcat 1080
cggcgacaac ttcttcgacg ggttgcaggt cgtgttcgga aacgtgctcg tgtggaagca 1140
ggtcatcagc cccacgcca tccgggtgca gacgcccccg cggcacatcc ccggggtggt 1200
ggaggtgacc ctctcctaca agtccaagca gttttgcaag ggatgccccg gccgctttgt 1260
ctacacagct ctgaacgagc ccaccattga ctacggattc cagaggctac agaaagtcat 1320
tcccagacac cccggagacc ccgagaggct gcccaaggaa gtgctgctga agcgggcggc 1380
cgacttgga gaagccctgt acggagtggc cggcagtaac caggagctgc tctgaagcg 1440
cgcgcgagac gtggccgagg ctctgtacag cacccccgcg gcacccgggc cgctcgcacc 1500
cctggccccg agccacccac actccgcctg cgtgggcatc aacgccttca gcagcccgct 1560
ggccatcgcc gtccgggacg ccaccccggg gcccgagccg ggtacgcgc gcagctgcag 1620
cagcgctgct ccccgcggtc tcgcgccag ccccggtctg cagcagagcg gctacggcg 1680
cggcctcgga gctggcctgg gcggctacgg cgcgccgggc gtggccggcc tcggcggtgc 1740
tggttcccc agcttctca atggtccac cgccacctcg cccttcgcca aggagcgct 1800
tcgccccgt gctgcgccc ccaagctccc caccacagcg ctgccccaga gccacggag 1860
aggggcttcc agaccagtct tttgaggatt ctgacaagtt ccactctcca gcccggggc 1920
ttcagggcct ggcatactcc taattacggt ctgcagctgt tcccatggag ccgagctgg 1980
aggtccctct gggattcaca gccacacccc ggatggtggc acagacagat gcagggccag 2040
ggccatgggc ggacctcaac ccgtgagctg aacggggaga ggccttcacc ccatgctcaa 2100
gcctccccgc tagcagcccc acaggcttct ctgcctccc tgtcttgggg tagtcagaag 2160
cccagcact gtgcagatgc tcttggcagg acagcatcgc agggaggtgc tgggattctg 2220
ggcctcactg tctgggtctt ggttccctctg aaagagatgg atcttgtgca gaccagggtt 2280
gttgagtga gggagcgtgg gatggggacc gtgggaaaga ggacagctca gggagaagt 2340
acctggaaag gtctgtttg catctgaccc atctcaactg gccagcatc ccaacttctc 2400
tgacgcgaaa ggggtggcgc ccgcagcctc gggaggcctg ccaggtctc cgtggagctt 2460
ccaacagctg cttggccccg cagctgcccc cacttcttt gagacctga ctctcatgct 2520
tgccgcatca tgctccctg tgggggcttt gggcatggag gaggcagaag agggggtgct 2580
aggcctctg tatttgggtt ctccccag tggtgtctc atggactctg gcccacaca 2640
ctcacaatga ctctggctgg cccacgcag cgggccagc cgccccccag gtggcctcac 2700
attctgctct gctaagttt gagaaaacag aacaataaac cagatgcagg tgggtcccg 2760
ccggcctctc acctgcctcc tt 2782

```

<210> 375

<211> 716

<212> DNA

<213> Homo sapiens

<300>

<308> NM_020188

<400> 375

```

acgggtggagc ggtggagggc gtcaactgggt ttccggctct ggcaagcggg tcagctgtct 60
gctccctagc agccggcctt cgggtcgggc gtctccgctg ctactgcgc ttcagttctc 120
ccggtgtggc cagcagtcgg gttgcactgc tgtgatccat cctcatctcc taaagatgca 180
tctgactta tctccacact tgcacactga agaatgcaac gtcttgatta acttgcttaa 240
ggaatgtcac aaaaatcaca acattctgaa attttttggg tattgtaatg atgttgatcg 300
ggagttgaga aaatgcctga agaatgagta cgtagaaaac aggaccaaga gcagggagca 360
tgccattgca atgcgaaaga aactttttta tctocagag gaatccgaaa aataaattgt 420
atcttctact gatgccttgg ctgagagaag acctaaagac tctgggttga tacctgaaag 480
aatctgtct tatttgggtt ccataatcct ttgaatggaa agtgacctgt gagagattga 540

```

accatggaga aatatgaaaa ccctggattc tgagtatttg ttgggcaggg cgtttagtac 600
 tgtctccctt ttaccagcaa acctgacttc accatgttta ttccctttgc ctacaaccag 660
 ttaatatctg agtaacttat ctcttcaat aaaataattt aaataaaaaa aaaaaa 716

<210> 376

<211> 3737

<212> DNA

<213> Homo sapiens

<300>

<308> NM_020974

<400> 376

ggcgtccgcg cacacctccc cgcgcgcgcg ccgccaccgc ccgcactccg ccgcctctgc 60
 ccgcaaccgc tgagccatcc atgggggtcg cgggcgcgcaa ccgtcccggg gcggcctggg 120
 cggtgctgct gctgtgctg ctgctgccgc cactgctgct gctggcgggg gccgtccgcg 180
 cgggtcgggg ccgtgccgcg gggcgcgcag aggatgtaga tgagtgtgcc caagggctag 240
 atgactgccg tgcgcagccc ctgtgtcaga acacacccac ctctacaag tgctcctgca 300
 agcctggcta ccaaggggaa ggcaggcagt gtgaggacat cgatgaatgt ggaaatgagc 360
 tcaatggagg ctgtgtccat gactgtttga atattccagg caattatcgt tgcacttgtt 420
 ttgatggcct catgttggct catgacgggc ataattgtct tgatgtggac gagtgcctgg 480
 agaacaatgg cggtgtccag catacctgtg tcaacgtcat ggggagctat gagtgtgct 540
 gcaaggaggg gtttttcctg agtgacaatc agcacacctg cattcacccg tcggaagagg 600
 gcoctgagct catgaataag gatcacggct gtatgcacat ctgcaaggag gccccaaggg 660
 gcagcgtcgc ctgtgagtgc aggcctgggt ttgagctggc caagaaccag agagactgca 720
 tottgacctg taaccatggg aacggtgggt gccagcactc ctgtgacgat acagccgatg 780
 gccagagtg cagctgccat ccacagtaca agatgcacac agatgggagg agctgccttg 840
 agcgagagga cactgtcctg gaggtgacag agagcaaac cacatcagtg gtggatgggg 900
 ataaacgggt gaaacggcgg ctgctcatgg aaactgtgct tgtcaacaat ggaggctgtg 960
 accgcacctg taaggatact tcgacagggt tccactgcag ttgtcctgtt ggattcactc 1020
 tccagtgtga tgggaagaca tgtaaagata ttgatgagtg ccagaccgcg aatggagggt 1080
 gtgatcattt ctgcaaaaac atcgtgggca gttttgactg cggctgcaag aaaggattta 1140
 aattattaac agatgagaag tcttgccaag atgtggatga gtgctctttg gataggacct 1200
 gtgaccacag ctgcataaac caccctggca catttgcttg tgcttgcaac cgagggtaca 1260
 ccctgtatgg cttacccac tgtggagaca ccaatgagtg cagcatcaac aacggaggct 1320
 gtcagcaggt ctgtgtgaac acagtgggca gctatgaatg ccagtgccac cctgggtaca 1380
 agctccactg gaataaaaaa gactgtgtgg aagtgaaggg gctcctgcc acaagtgtgt 1440
 caccocgtgt gtccctgcac tgcggtgaag gtgggtggag agacgggtgc ttcctcagat 1500
 gtcactctgg cattcacctc tcttcagatg tcaccacat caggacaagt gtaaccttta 1560
 agctaaatga aggcaagtgt agtttgaaaa atgctgagct gtttcccgag ggtctgcgac 1620
 cagcactacc agagaagcac agctcagtaa aagagagctt ccgctacgta aaccttacat 1680
 gcagctctgg caagcaagtc ccaggagccc ctggccgacc aagcaccctt aaggaaatgt 1740
 ttatcactgt tgagtttgag cttgaaacta accaaaagga ggtgacagct tcttgtgacc 1800
 tgagctgcat cgtaaagcga accgagaagc ggctccgtaa agccatccgc acgctcagaa 1860
 aggcgtcca caggagcag tttcacctcc agctctcagg catgaacctc gacgtggcta 1920
 aaaagcctcc cagaacatct gaacgccagg cagagtcctg tggagtgggc cagggctcatg 1980
 cagaaaacca atgtgtcagt tgcagggtgt ggacctatta tgatggagca cgagaacgct 2040
 gcattttatg tccaaatgga accttccaaa atgaggaagg acaaatgact tgtgaacat 2100
 gcccaagacc aggaattctt ggggccctga agacccaga agcttggat atgtctgat 2160
 gtggaggtct gtgtcaacct ggtgaatatt ctgcagatgg ctttgcacct tgccagctct 2220
 gtgccttggg cacgttccag cctgaagctg ctgcgaactt ctgcttcccc tgtggaggag 2280
 gccttggcac caaacatcag ggagctactt cctttcagga ctgtgaaacc agagttcaat 2340
 gttcacctgg acatttctac aacaccacca ctacccgatg tattcgttgc ccagtgggaa 2400
 cataccagcc tgaatttgga aaaaataatt gtgtttcttg ccagggaaat actacgactg 2460
 actttgatgg ctccacaaac ataaccagt gtaaaaacag aagatgtgga ggggagctgg 2520
 gagatttcac tgggtacatt gaatccccaa actaccagg caattacca gccaacaccg 2580
 agtgtacgtg gaccatcaac ccacccccca agcgccgcat cctgatcgtg gtccttgaga 2640
 tcttcttgc ctatagggag gactgtgggg actatctggt gatgcggaaa acctctcat 2700
 ccaattctgt gaaacctgac gaaacctgcc agacctacg acgccccat gccttcacct 2760
 ccagggtcaa gaagctgttg attcagttca agtccaatga agggaacagc gctagagggt 2820
 tccagggtccc atacgtgaca tatgatgagg actaccagga actcattgaa gacatagttc 2880

```

gagatggcag gctctatgca tctgagaacc atcaggaaat acttaaggat aagaaactta 2940
tcaaggctct gtttgatgtc ctggccatc cccagaacta tttcaagtac acagcccagg 3000
agtcccgaga gatgtttcca agatcgttca tccgattgct acgttccaaa gtgtccagggt 3060
ttttgagacc ttacaaatga ctcagcccac gtgccactca atacaaatgt tctgctatag 3120
ggttggtggg acagagctgt cttccttctg catgtcagca cagtcgggta ttgctgcctc 3180
ccgtatcagt gactcattag agttcaattt ttatagataa tacagatatt ttggtaaatt 3240
gaacttggtt tttctttccc agcatcgtgg atgtagactg agaatggctt tgagtggcat 3300
cagcttctca ctgctgtggg cggatgtctt ggatagatca cgggctggct gagctggact 3360
ttggtcagcc taggtgagac tcacctgtcc ttctggggtc ttactcctcc tcaaggagtc 3420
tgtagtggaa aggagccac agaataagct gcttattctg aaacttcagc ttcctctagc 3480
ccggccctct ctaaggagc cctctgcact cgtgtgcagg ctctgaccag gcagaacagg 3540
caagagggga gggaaggaga cccctgcagg ctccctccac ccaccttgag acctgggagg 3600
actcagtttc tccacagcct tctccagcct gtgtgataca agtttgatcc caggaaacttg 3660
agttctaagc agtgctcgtg aaaaaaaaaa gcagaaagaa ttagaaataa ataaaaacta 3720
agcacttctg gagacat 3737

```

<210> 377

<211> 3304

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000127

<400> 377

```

gcgaccgaac gcggcggtcg gcagcgttcg cgcggggggc tgcgaagcgc tgctcggggc 60
cggcactgcc gcgggggagg acgcgcgcgc gccgccacc agcgcgcgcg ccgccgcgcg 120
ctccagcccg gccgcgcgcg gtcccggggg cggcccccgc gagcgcagga gtaaaccaccg 180
ccggagtctt ggagccgctg cagaagggaa taaagagaga tgcagggatt tgtgagggtta 240
cggcgcccca gctgcaagat gcaactagccg gctgaacccg ggatcggctg acttggttga 300
accggagtgc tctgcacgga gagtgggtga tgagtgaag ttgccttccc ggggctcatt 360
ttccacgctg ccgagaggaa tccgagaggg aaggcaatca cttcgtcttg ccattgattg 420
ggatatcgga gctttttttt tctccctct ctctttcttt tctcctgtct tgttgcatgc 480
aagaaaatta cagtccgctg ctgcgccgc ctgggtgcga gatattcagc cccgctctct 540
ccgtgcatc gtgcaaccca aagatgaaag accgaagggg agaaagttaa agaaatcgcc 600
cacatgcgct ggatcagtc acggtctggg gaaaggcatc cagagaagggt gggagcggag 660
agtttgaagt ctttacaggc gggaagatgg cggactggag ctgaaagtgt tgattgggaa 720
acttgggtga ttcttgtgtt tatttacaat cctcttgacc caggcaggac acatgcaggc 780
caaaaaacgc tatttcatcc tgctctcagc tggctcttgt ctgcctttt tgttttatt 840
cggaggcttg cagtttaggg catcgaggag ccacagccgg agagaagaac acagcggtag 900
gaatggcttg caccacccca gtccggatca tttctggccc cgcttcccgg agcctctgcg 960
ccccttcgtt ccttgggatc aattggaaaa cgaggattcc agcgtgcaca tttccccccg 1020
gcagaagcga gatgccaact ccagcatcta caaaggcaag aagtgcgcga tggagtcctg 1080
cttcgatttc accctttgca agaaaaacgg cttcaaagtc tacgtatacc cacagcaaaa 1140
aggggagaaa atcgccgaaa gttaccaaaa cattctagcg gccatcgagg gctccaggtt 1200
ctacacctcg gaccccagcc agcgctgcct ctttgtctg agtctggata ctttagacag 1260
agaccagttg tcacctcagt atgtgcacaa tttgagatcc aaagtgcaga gtctccactt 1320
gtggaacaat ggtaggaatc atttaatttt taatttata tccggcactt ggcctgacta 1380
caccgaggac gtggggtttg acatcgcca ggcgatgctg gccaaagcca gcatcagtac 1440
tgaaaacttc cgaccaact ttgatgtttc tattcccctc ttttctaagg atcatcccag 1500
gacaggaggg gagagggggt ttttgaagtt caacaccatc cctcctctca ggaagtacat 1560
gctggtattc aaggggaaga ggtacctgac agggatagga tcagacacca ggaatgcctt 1620
atatcacgct cataacgggg aggacgttgt gctcctcacc acctgcaagc atggcaaga 1680
ctggcaaaaag cacaaggatt ctgcgtgtga cacagacaac accgagtatg agaagtatga 1740
ttatcgggaa atgctgcaca atgccacttt ctgtctgggt cctcgtggtc gcaggcttgg 1800
gtccttcaga ttcttgagg ctttcagggc tgcctgcgtc cctgtgatgc tcagcaatgg 1860
atgggagttg ccattctctg aagtatttaa ttggaaccaa gctgccgtca taggcgatga 1920
gagattgtta ttacagattc cttctacaat caggtctatt catcaggata aaatcctagc 1980
acttagacag cagacacaat tcttggtgga ggcttatttt tcttcagttg agaagattgt 2040
attacacata ctacagatta ttcaggacag aatattcaag cacatatcac gtaacagttt 2100
aatatggaac aaacatcctg gaggattgtt cgtactacca cagtattcat cttatctggg 2160

```

```

agattttcct tactactatg ctaatttagg tttaaagccc ccctccaaat tcaactgcagt 2220
catccatgag gtgaccccc ttggtctctca gtcccagcca gtgttgaagc ttctcgtggc 2280
tgacagccaag tcccagttact gtgccagat catagtctta tgggaattgtg acaagcccct 2340
accagccaaa caccgctggc ctgccactgc tgtgcctgtc gtcgtcattg aaggagagag 2400
caaggttatg agcagccgtt ttctgcccta cgacaacatc atcacagacg ccgtgctcag 2460
ccttgacgag gacacggtgc tttcaacaac agaggtggat ttgccttca cagtgtggca 2520
gagcttcctt gagaggattg tggggtaccc cgcgcgcagc cactttctggg ataactctaa 2580
ggagcgggtg ggatacacat caaagtggac gaacgactac tccatggtgt tgacaggagc 2640
tgctatttac cacaaatatt atcactacct atactcccat tacctgccag ccagcctgaa 2700
gaacatggtg gaccaattgg ccaattgtga ggacattctc atgaacttcc tgggtgtctgc 2760
tgtgacaaaa ttgcctccaa tcaaagtgc ccagaagaag cagtataagg agacaatgat 2820
gggacagact tctcgggctt cccgttgggc tgaccctgac cactttgccc agcgacagag 2880
ctgcatgaat acgtttgccg gctggtttgg ctacatgccg ctgatccact ctccagatgag 2940
gctcgacccc gtccctctta aagaccaggt ctctattttg aggaagaaat accgagacat 3000
tgagcgactt tgaggaatcc ggctgagtgg gggaggggaa gcaagaaggg atgggggtca 3060
agctgctctc tcttcccagt gcagatccac tcatcagcag agccagattg tgccaaactat 3120
ccaaaaactt agatgagcag aatgacaaaa aaaaaaaagg ccaatgagaa ctcaactcct 3180
ggctcctggg actgctccaa actcacctca ctggcttctg tgtcccaaga 3240
ctaggttggg acagtttaat tatggaacat taaataatta tttttgaaaa aaaaaaaaaa 3300
aaaa 3304

```

<210> 378

<211> 7680

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002019

<400> 378

```

gcgacactc ctctcggctc ctccccggca gggcgggcgg ctccggagcgg gctccggggc 60
tcgggtgcag cggccagcgg gcctggcggc gaggattacc cggggaagtg gttgtctcct 120
ggctggagcc gcgagacggg gcgtcagggc cgggggcccgg cggcggcgaa cgagaggacg 180
gactctggcg gccgggtcgt tggcgggggg agcgcgggca ccgggcgagc aggcgcgcgtc 240
gcgctcacca tggtcagcta ctgggacacc ggggtcctgc tgtgcgcgct gctcagctgt 300
ctgcttctca caggatctag ttcaggttca aaattaaaag atcctgaact gagtttaaaa 360
ggcaccacgc acatcatgca agcaggccag acactgcatc tccaatgcag gggggaagca 420
gcccataaat ggtctttgccc tgaaatgggt agtaaggaaa gcgaaaggct gagcataact 480
aaatctgcct gtggaagaaa tggcaacaaa ttctgcagta ctttaacctt gaacacagct 540
caagcaaaccc acactggctt ctacagctgc aaatatctag ctgtacctac ttcaaagaag 600
aaggaaacag aatctgcaat ctatatattt attagtgata caggtagacc tttcgtagag 660
atgtacagtg aaatccccga aattatacac atgactgaag gaaggagct cgtcattccc 720
tgccgggtta cgtcacctaa catcactgtt actttaaaaa agtttccact tgacactttg 780
atccctgatg gaaaacgcat aatctgggac agtagaaaag gcttcatcat atcaaatgca 840
acgtacaaag aaatagggct tctgacctgt gaagcaacag tcaatgggca tttgtataag 900
acaaactatc tcacacatcg acaaaccaat attgtactgc taccactccc 1020
cgcccagtc aattacttag aggccatact cttgtcctca attgtactgc taccactccc 1080
ttgaacacga gagttcaaat gacctggagt taccctgatg aaaaaataa gagagcttcc 1140
gtaaggcgac gaattgacca aagcaattcc catgccaaac tattctacag tgttcttact 1200
attgacaaaa tgcagaacaa agacaaagga ctttatactt gtcgtgtaag gagtggacca 1260
tcattcaaat ctgttaacac ctacgtgcat atatatgata aagcattcat cactgtgaaa 1320
catcgaaaac agcagggtgt tgaaaccgta gctggcaagc ggtcttaccg gctctctatg 1380
aaagtgaagg caattccctc gccggaagtt gtaggttaa aagatgggtt acctgcgact 1440
gagaaatctg ctgcgtatct gactcgtggc tactcgttaa ttatcaagga cgttaactgaa 1500
gaggatgcag ggaattatac aatcttgctg agcataaaac agtcaaatgt gtttaaaaac 1560
ctcactgcc aactaattgt caatgtgaaa cccagattt acgaaaaggc cgtgtcatcg 1620
tttccagacc cggctctcta cccactgggc agcagacaaa tccgtacttg taccgcata 1680
ggtatccctc aacctacaat caagtgggtc tggcaccctt gtaaccataa tcattccgaa 1740
gcaagggtgt acttttggtc caataatgaa gagtccctta tcctggatgc tgacagcaac 1800
atgggaacaa gaattgagag catcactcag cgcattggca taatagaagg aaagaataag 1860
atggctagca ccttgggtgt ggctgactct agaatttctg gaattctacat ttgcatagct

```

tccaataaag	ttgggactgt	gggaagaaac	ataagctttt	atatcacaga	tgtgccaaat	1920
gggtttcatg	ttaacttgga	aaaaatgccg	acggaaggag	aggacctgaa	actgtcttgc	1980
acagttaaca	agttcttata	cagagacgtt	acttggtatt	tactgcggac	agttaataac	2040
agaacaatgc	actacagtat	tagcaagcaa	aaaatggcca	tcactaagga	gcactccatc	2100
actcttaatc	ttaccatcat	gaatgtttcc	ctgcaagatt	caggcaccta	tgcttcgaga	2160
gccaggaatg	tatacacagg	ggaagaaatc	ctccagaaga	aagaaattac	aatcagagat	2220
caggaagcac	catacctcct	gcgaaacctc	agtgatcaca	cagtggccat	cagcagttcc	2280
accactttag	actgtcatgc	taatgggtgc	cccgagcctc	agatcacttg	gtttaaaaac	2340
aaccacaata	tacaacaaga	gcctggaatt	attttaggac	caggaagcag	cacgtgtgtt	2400
attgaaagag	tcacagaaga	ggatgaaggt	gtctatcact	gcaaagccac	caaccagaag	2460
ggctctgtgg	aaagttcagc	atacctcact	gttcaaggaa	cctcggaaca	gtctaattctg	2520
gagctgatca	ctctaactatg	cacctgtgtg	gctgcgactc	tcttctggct	cctattaacc	2580
ctccttatcc	gaaaaatgaa	aaggtcttct	tctgaaataa	agactgacta	cctatcaatt	2640
ataatggacc	cagatgaagt	tcctttggat	gagcagtgtg	agcggctccc	ttatgatgcc	2700
agcaagtggg	agtttgcccg	ggagagactt	aaactgggca	aatcacttgg	aagaggggct	2760
tttgaaaaag	tggttcaagc	atcagcattt	ggcattaaga	aatcacctac	gtgccggact	2820
gtggctgtga	aaatgctgaa	agagggggcc	acggccagcg	agtacaaagc	tctgatgact	2880
gagctaaaaa	tcttgacca	cattggccac	catctgaacg	tggttaacct	gctgggagcc	2940
tgaccaagc	aaggagggcc	tctgatgggtg	attgttgaat	actgcaaata	tggaatctc	3000
tccaactacc	tcaagagcaa	acgtgactta	ttttttctca	acaaggatgc	agcactacac	3060
atggagccta	agaaagaaaa	aatggagcca	ggcctggaac	aaggcaagaa	accaagacta	3120
gatagcgtca	ccagcagcga	aagctttgcg	agctccggct	ttcaggaaga	taaaagtctg	3180
agtgtgttg	aggaagagga	ggattctgac	ggtttctaca	aggagcccat	cactatggaa	3240
gatctgattt	cttacagttt	tcaagtggcc	agaggcatgg	agttcctgtc	ttccagaaag	3300
tgcatctcatc	gggacctggc	agcgagaaac	attcttttat	ctgagaacaa	cgtgggtgaag	3360
atttgtgatt	ttggccttgc	ccgggatatt	tataagaacc	ccgattatgt	gagaaaagga	3420
gatactcgac	ttcctctgaa	atggatggct	ccgaatcta	tctttgacaa	aatctacagc	3480
accaagagcg	acgtgtgtgc	ttacggagta	ttgctgtggg	aaatcttctc	cttaggtggg	3540
tctccatacc	caggagtaca	aatggatgag	gacttttgca	gtcgccctgag	ggaaggcatg	3600
aggatgagag	ctcctgagta	ctctactcct	gaaatctatc	agatcatgct	ggactgctgg	3660
cacagagacc	caaaagaaag	gccaagattt	gcagaacttg	tggaataact	aggtgatttg	3720
cttcaagcaa	atgtacaaca	ggatggtaaa	gactacatcc	caatcaatgc	catactgaca	3780
ggaaatagtg	ggttttacata	ctcaactcct	gccttctctg	aggacttctt	caaggaaagt	3840
atttcagctc	cgaagtttaa	ttcaggaagc	tctgatgatg	tcagatatgt	aaatgctttc	3900
aagttcatga	gcctggaaag	aatcaaaaacc	tttgaaagaac	ttttaccgaa	tgccacctcc	3960
atgtttgatg	actaccaggg	cgacagcagc	actctgttgg	cctctcccat	gctgaagcgc	4020
ttcacctgga	ctgacagcaa	acccaaggcc	tcgctcaaga	ttgacttgag	agtaaccagt	4080
aaaagtaagg	agtcggggct	gtctgatgtc	agcaggccca	gtttctgcca	ttccagctgt	4140
gggcacgtca	gcgaaggcaa	gcgcaggttc	acctaagacc	acgtgagct	ggaaaggaaa	4200
atcgctgtct	gctccccgcc	cccagactac	aactcgggtg	tcctgtactc	cacccccacc	4260
atctagagtt	tgacacgaag	ccttatttct	agaagcacat	gtgtattttat	acccccagga	4320
aactagcttt	tgccagtatt	atgcataat	aagtttacac	ctttatcttt	ccatgggagc	4380
cagctgcttt	ttgtgatttt	tttaatatgt	cttttttttt	ttgactaaca	agaatgtaac	4440
tccagataga	gaaatagtga	caagtgaaga	acactactgc	taaactcctca	tgttactcag	4500
tgttagagaa	atccttccta	aacccaatga	cttccctgct	ccaacccccg	ccacctcagg	4560
gcacgcagga	ccagtttgat	tgaggagctg	cactgatcac	ccaatgcac	acgtacccca	4620
ctggggcagc	cctgcagccc	aaaacccagg	gcaacaagcc	cgtagacccc	aggggatcac	4680
tggtgtgctt	gagcaacatc	tcgggagtc	tctagcaggc	ctaagacatg	tgaggaggaa	4740
aaggaaaaaa	agcaaaaagc	aaggagagaa	agagaaaccg	ggagaaggca	tgagaaagaa	4800
tttgagacgc	accatgtggg	cacggagggg	gacggggctc	agcaatgcc	tttcagtggc	4860
ttccagctc	tgaccttct	acatttgagg	gcccagccag	gagcagatgg	acagcgatga	4920
ggggacattt	tctggttct	gggaggcaag	aaaaggacaa	atatcttttt	tggaactaaa	4980
gcaaatttta	gacctttacc	tatggaagtg	gttctatgtc	cattctcatt	cgtggcatgt	5040
tttgatttgt	agcactgagg	gtggcactca	actctgagcc	catacttttg	gctcctctag	5100
taagatgoac	tgaaaactta	gccagagtta	ggttgtctcc	aggccatgat	ggccttacac	5160
tgaaaatgtc	acattctatt	ttgggtatta	atatatagtc	cagacactta	actcaatttc	5220
ttggtattat	tctgttttgc	acagttagtt	gtgaaagaaa	gctgagaaga	atgaaaatgc	5280
agtctgagg	agagttttct	ccatatcaaa	acgagggctg	atggaggaaa	aaggtcaata	5340
aggtcaagg	aagacccgt	ctctatacca	accaaaccaa	ttcaccaaca	cagttgggac	5400
ccaaaacaca	ggaagtcatg	caogtttct	tttcaattta	tggggattcc	actatctcac	5460
actaatctga	aaggatgtgg	aagagcatta	gctggcgcat	attaagcact	ttaagctcct	5520


```

tgagtaaaaa ggtggtatgt aatttatgca aggtatttct ccagttggga ctcaggatat 5580
tagttaatga gccatcacta gaagaaaagc ccattttcaa ctgotttgaa acttgccctgg 5640
ggtctgagca tgatgggaat agggagacag ggtaggaaaag ggcgcctact cttcagggtc 5700
taaagatcaa gtgggccttg gatcgctaag ctggctctgt ttgatgctat ttatgcaagt 5760
tagggctctat gtatttagga tgccgcctact cttcagggtc taaagatcaa gtgggccttg 5820
gatcgctaag ctggctctgt ttgatgctat ttatgcaagt tagggctctat gtatttagga 5880
tgtctgcacc ttctgcagcc agtcagaagc tggagaggca acagtggatt gctgcttctt 5940
ggggagaaga gtatgcttcc ttttatccat gtaatttaac tgtagaacct gagctctaag 6000
taaccgaaga atgtatgcct ctgttcttat gtgccacatc cttgttttaa ggctctctgt 6060
atgaagagat gggaccgtca tcagcacatt ccctagttag cctactggct cctggcagcg 6120
gcttttgtgg aagactcact agccagaaga gaggagtggg acagtccctc ccaccaagat 6180
ctaaatccaa acaaaagcag gctagagcca gaagagagga caaatctttg ttgttccctc 6240
tctttacaca tacgcaaacc acctgtgaca gctggcaatt ttataaatca ggtaactgga 6300
aggaggttaa actcagaaaa aagaagacct cagtcaattc tctacttttt tttttttttt 6360
tccaaatcag ataatagccc agcaaatagt gtaacaaat aaaaccttag ctgttcatgt 6420
cttgatttca ataattaatt cttaatcatt aagagaccat aataaatact ctttttcaag 6480
agaaaagcaa aaccattaga attgttactc agctccttca aactcaggtt tgtagcatac 6540
atgagtccat ccatcagtca aagaatgggt ccactctggg tcttaatgta gaaagaaaaa 6600
tgagagacttg taataatgag ctagttaaaa agtgcttggt cattaaaaata gcactgaaaa 6660
ttgaaacatg aattaactga taatattcca atcatttgcc atttatgaca aaaatgggtg 6720
gcactaacia agaacgagca cttcctttca gagtttctga gataatgtac gtggaacagt 6780
ctgggtggaa tggggctgaa accatgtgca agtctgtgtc ttgtcagtcc aagaagtac 6840
accgagatgt taatttttagg gaccctgccc ttgtttccta gccacaaga atgcaaacat 6900
caaacagata ctgcctagcc tcatttaaat tgattaaagg aggagtgcac ctttggccga 6960
cagtgggtgta actgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tgtgggtgtg 7020
ggtgtatgtg tgttttgtgc ataactattt aaggaaactg gaatttttaa gttactttta 7080
tacaaccaa gaatatatgc tacagatata agacagacat ggtttggtcc tatatttcta 7140
gtcatgatga atgtattttg tataccatct tcatataata tacttaaaaa tatttcttaa 7200
ttgggatttg taatcgatcc aacttaattg ataaacttgg caactgcttt tatgttctgt 7260
ctccttccat aaatttttca aaataacta tcaacaaaga aaaagctctt ttttttcta 7320
aaataaactc aaatttatcc ttgttttagg cagagaaaaa ttaagaaaaa ctttgaaatg 7380
gtctcaaaaa attgctaaat attttcaatg gaaaactaaa tgtagtttta gctgattgta 7440
tggggttttc gaacctttca ctttttgttt gttttaccta tttcacaact gtgtaaattg 7500
ccaataatc ctgtccatga aaatgcaaat tatccagtgt agatatattt gaccatcacc 7560
ctatggatat tggctagttt tgccctttatt aagcaaatc atttcagcct gaatgtctgc 7620
ctatatattc tctgctcttt gtatttctct ttgaaccctg taaaacatcc tgtggcactc 7680

```

<210> 379

<211> 2679

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002073

<400> 379

```

gagaccagga ccatgggatg tcggcaaagc tcagaggaaa aagaagcagc ccggcggtcc 60
cggagaattg accgccacct gcgctcagag agccagcggc aacgccgga aatcaagctg 120
ctcctgctgg gcaccagcaa ctgagcaag agcaccatcg tcaaacagat gaagatcatc 180
cacagcggcg gcttcaacct ggaggcctgc aaggagtaca agcccctcat catctacaat 240
gccatcgact cgctgaccog catcatccgg gccctggcgg cctcaggat cgacttccac 300
aaccocgacc gcgctaocga cgctgtgcag ctctttgcgc tgacggggcc cgctgagagc 360
aaggcgagga tcacaccgga gctgctgggt gtcatgcgac ggctctgggc cgaccaggg 420
gcacaggcct gcttcagccg ctccagcgag taccacctgg aggacaacgc ggcctactac 480
ctgaacgacc tggagcgcat cgccgcagct gactatatcc cactgtcga ggacatcctg 540
cgctcccggg acatgaccac gggcattgtg gagaacaagt tcaccttcaa ggagctcacc 600
ttcaagatgg tggacgtggg ggggcagagg tcagagcgca aaaagtggat cactgcttc 660
gagggcgta cagccatcat cttctgtgtg gagctcagcg gctacgacct gaaactctac 720

```

gaggataacc	agacaagtgc	gatggcagag	agcttgccgc	tctttgactc	catctgcaac	780
aacaactggt	tcatcaacac	ctcactcatc	ctcttcctga	acaagaagga	cctgctggca	840
gagaagatcc	gccgcatccc	gtccaccatc	tgctttcccg	agtacaaggg	ccagaacacg	900
tacgaggagg	ccgctgtcta	catccagcgg	cagtttgaag	acctgaaccg	caacaaggag	960
accaaggaga	tctactccca	cttcacctgc	gccaccgaca	ccagtaacat	ccagtttgtc	1020
ttcgacgcgg	tgacagacgt	catcatacag	aacaatctca	agtacattgg	cctttgctga	1080
ggagctgggc	ccggggcgcc	tgccataggt	gaaacccacg	gggtgtcatg	ccccaacgcg	1140
tgctagagag	gccaatcca	ggggcagaaa	acagggggcc	taaagaatgt	cccccacccc	1200
ttggcctctg	cctccttggc	cccacatttc	tgcaaacata	aatattttacg	gatagattgc	1260
taggtagata	gacacacaca	catgcacaca	cacacatctg	gagatggcaa	aatcctctaa	1320
aatgtcgagg	tctcttgaag	acttgagaag	ctgtcacaag	gtcactacaa	gccaacctg	1380
ccccttcaact	ttgccttcct	gagttggccc	cactccactt	gggggtctgc	attggattgt	1440
tagggatagg	cagcaggcgt	gaggcaaggt	aggccaactg	cacccctgtc	acctggaggga	1500
gggcccgtc	gctgcccgag	ctctggccta	gggaccttgc	cgctgaccaa	gagggaggac	1560
cagtgcaggg	tctgtgcacc	ttccctgctg	gcctgcacac	agctgctcag	caccatttca	1620
ttctggacct	gggaccttag	gagccgggtg	acagcactaa	ccagacctcc	agccactcac	1680
agctcttttt	aaaaaacagc	ttcaaaatat	gcagcaaaaa	ccaatacaac	aaaacgagtg	1740
gcacgattta	tttcaaaacta	ggccagctgg	gattccagct	tttctcttac	tagtctgatg	1800
ttttataaat	caaaacctgg	ttttccttct	ctggcatttt	tttttgtttt	ttgttttttg	1860
gttttttttt	tttttttggc	caaatctcgt	ggtgtttcgc	agaaaaaaat	ccagaaaatt	1920
tcaaatgcag	ttgagtattc	tttttttaaat	gcagattttc	aaaacatatt	ttttttcagg	1980
tggtcttttt	tgtgtctggc	ttgctgagtg	taaaagttgt	tatctggacg	atctgtctct	2040
ctgctccaaa	gaaatttttg	agtgagtggc	agtcctgcgc	cagcctcgcg	ggacacgtgt	2100
tgacataaag	cctctgcagt	gtcctcttgt	taatgggtgg	gttttctgct	ttgtttttat	2160
ttaagaaaat	aaacacgaca	tattttaaaga	aggttctttc	acctgggagc	aatgaacaa	2220
tagctaagtg	tcttggtatt	taaagagtaa	attattttgtg	gctttgctga	gtgaagggaag	2280
gggagcaagg	ggtggtgccc	ctgggtccag	catgccccgc	gcctgagact	ggctggaaat	2340
gctctgactc	ctgtgaaggc	acagccagcg	ttgtggcctg	agggaggccc	tgctgggacc	2400
ctgatctggg	ccttctctgtc	ccagggccta	tgggcaactg	cgttgaaagg	acgttcgcca	2460
agggccgtgt	gtaaatacga	actgcgccat	ggagaggaga	ggcactgccg	gagcccttgc	2520
cagatctccc	tccctctctc	tgtgcagtag	ctgtgtgtcc	gaggtcagtg	tgcggaatca	2580
cagccaagga	cgtgaagaga	tgtacggggg	aaagagaagc	tggggatttg	atgaaagtca	2640
aaggttgtct	actttaagaa	aataaaatac	cctgaatgg			2679

<210> 380

<211> 3337

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000436

<400> 380

gtcgagcctc	tagcccgccc	gggtttcctt	cgcagtcgcg	caccgacgct	caaacgcgcg	60
ctccaaccgg	cagcctcctc	ctgcctcacc	gcccgaagat	ggcggtctct	aaactcctct	120
cctccgggct	tcggctctgc	gcctctgccc	gcggatctgg	ggcaacctgg	tacaagggat	180
gtgtttgttc	cttttccaac	agtgtctatc	gccataccaa	gttttatata	gatccagtag	240
aagctgtaaa	agacatccct	gatggtgcc	cggttttggg	tggtgggttt	gggctatgtg	300
gaattccaga	gaatcttata	gatgctttac	tgaaaactgg	agtaaaagga	ctaactgcag	360
tcagcaacaa	tgacggggtt	gacaattttg	gtttggggct	tttgcttcgg	tcgaagcaga	420
taaaacgc	ggtctcttca	tatgtgggag	aaaatgcaga	atltgaacga	cagtacttat	480
ctgggtgaatt	agaagtggag	ctgacaccac	agggcacact	tgacagagag	atccgtgcag	540
gcggggctgg	agttctctga	ttttacaccc	caacagggta	tgggaccctg	gtacaagaag	600
gaggatcgcc	catcaaatac	aacaaagatg	gcagtgttgc	cattgccagt	aagccaagag	660
aggtgaggga	gttcaatggt	cagcacttta	ttttggagga	agcaattaca	ggggattttg	720
ctttgtgtaa	agcctggaag	gcggaccgag	caggaaacgt	gatttttcagg	aaaagtgc	780
ggaatttcaa	cttgccaatg	tgcaaagctg	cagaaaccac	agtggttagag	gttgaagaaa	840
ttgtggatat	tgagacattt	gctccagaag	acatccatat	tcctcagatt	tatgtacatc	900
gccttataaa	gggagaaaaa	tatgagaaaa	gaattgagcg	tttatcaatc	cggaagagg	960
gagatgggga	agccaatct	gctaaccctg	gagatgacgt	aagggaacga	atcatcaaga	1020
gggcccgtct	tgagtttgag	gatggcatgt	atgctaattt	gggcatagga	atccctctcc	1080

tgccagcaa	ttttatcagc	ccaaatataa	ctgttcacat	tcaaagtga	aatggagttc	1140
tggtttggg	tccatatcca	cgacaacatg	aagctgatgc	agatctcatc	aatgcaggca	1200
aggaaacagt	tactattcct	ccaggagcct	cttttttctc	cagcgatgaa	tcatttgcaa	1260
tgattagagg	tggaacagtc	gatctgacaa	tgctaggagc	gatgcagggt	tccaaatatg	1320
gtgacctggc	taactggatg	atacctggga	agatgggtga	aggaatggga	ggtgctatgg	1380
athtagtgtc	cagtgcgaaa	accaaagtgg	tggtcaccat	ggagcattct	gcaaagggaa	1440
atgcacataa	aatcatggag	aaatgtacat	taccattgac	tggaagcaa	tgtgtcaacc	1500
gcattattac	tgaaaaggct	gtgtttgatg	tggaacaaga	gaaagggttg	actctgattg	1560
agctctggga	aggcctgaca	gtggatgacg	tacaaaagag	tactgggtgt	gattttgcag	1620
tttcaccaa	actcatgcc	atgcagcaga	tcgcaaattg	aaatatggat	atgtgtacca	1680
ggctgcgtgt	ttttcatttt	aaacacacaa	gatttaattg	aaaggacatc	aataatcata	1740
atgtgtatt	taacaggtgg	ttttttatta	gttttcttgt	gtttcagact	ttatgcagcc	1800
atataaactg	ttctctaggc	atgctgtgac	attttaataa	aaagcaaaag	gagcatttat	1860
aattatctca	tttgtaaagg	ctgagaaggt	tggttttata	ataggtaatt	atattgaatg	1920
cattttcaat	gaatatggta	tgtatgctaa	attatatgaa	cctttcccca	agaagggccc	1980
tagaaattga	tgtggctttc	ctcttaataa	tttaattatta	gtcctgaaag	aaagataaca	2040
tatgtgattt	ttgtggttag	gagagttgct	gtcatgattg	ttttttcttc	agcctcctct	2100
gacttttctt	ttggggcttc	agatttttatg	attacatctt	gtccccctag	aacatccccc	2160
ttcctcccat	actgctttta	aacagatgcc	caagaaggca	agcaggaatg	cctcttgttg	2220
gggagggcag	ggagaaataa	ctagttcaaa	ccaactatct	atctatgctt	tgcaaagact	2280
aaggcgtatt	ataggaagag	ggctagaaac	ctaactgatt	cttctcagtt	ttctcatttt	2340
aaaacagccc	agtattcctt	tgtatcctca	agggctcctg	agaatacttc	tgttattgaa	2400
accctgtggg	ctactgtgac	tgtacctcct	ctcaagccaa	gaagggctgt	gggataattt	2460
accatgaatc	cttagtagca	atgacagcag	agttaaaaaa	taaaagggtg	tttactttca	2520
ggctcttggt	ttgggtcaga	ggagatttta	aatattgaat	gacacttcta	cagaacaacg	2580
gtttttcttc	tgccaaggct	acttccttta	acgaagtgcc	tttaattcag	ccttatccaa	2640
ctagggaaaa	taatgttgga	caagtctagg	atgtgaagag	tcagtgaact	tttagtgtca	2700
gggaataaac	atgggtgggt	gatttaggtt	gaaaaaaact	tccttagagg	tattttattct	2760
caatacctga	caggggcccc	tggaatgac	ttcagaagca	tcccggataa	tagatgggta	2820
aaaagtctag	gcaccctgaa	gaacaggtga	gacagctggc	ctctggacag	aggtaggcat	2880
agtagcagta	gatatatcat	tcctctggtc	ctaaatatac	aaacttattc	atgtttttag	2940
gtgatgatgg	tcattgaaac	tcacttcttt	tcagggtgtag	ctacaattgt	gtaatgtaca	3000
atattagaga	aaggacaggc	tttttatgag	taacacacac	catatataaa	acagcctttc	3060
tggtcgacca	catggttaaa	tgcatacctt	ccagtagctg	gggggaaaat	gacccttctt	3120
agaatgtgca	agttccatag	agtaatatat	tgatatgatt	ttgaaaagaa	ttgttgatag	3180
ttacatcttc	aaacttatca	ttccagtatg	catctttaag	ataatgtgat	tctaagtaga	3240
tgactttata	ttcttgatta	aagagtgcata	tacatgttaa	gaaatgcatt	aaggaataca	3300
ataaatattc	taaactgatg	aaaaaaaaaa	aaaaaaa			3337

<210> 381

<211> 1020

<212> DNA

<213> Homo sapiens

<300>

<308> Contig55377_RC

<400> 381

tagttttcaa	aaaaatattg	atctggatca	gatgggctga	taaggcccaa	acacttataa	60
accaaagctg	caggctctca	aggtcttcag	agaagctaca	caaaatacta	tgattttcatt	120
agcatccaca	tatgtatcat	tttccacaac	attttacaat	tctcttggtt	cacagggaag	180
gtgaactaat	tataatgaaa	tatcagtaat	gtatcttaaa	taagtaacat	gttgctggaa	240
aaaatttaag	ttaaaatgca	ctacataata	gtgcagaaac	aaagtgcagg	tggcaatttt	300
aagacatgaa	aacaatgctg	aatcagtagt	gtttggagtg	atacattttg	cccatttttg	360
gccagtccca	gaacagtact	caccacaagc	tgcatgccaa	atatgaacta	atgaggctaa	420
aacactgaag	accaaataga	accactctag	agtctgtttc	acaatgagat	actcctcccc	480
gttctggaaa	cagcgatctg	ctgtagttcc	tatcctcctg	tggcacctat	ccctgctaca	540
gtacctataa	ccagatgtcc	tccatcatcc	agcctccagg	acaagtgggt	ctcagtgaag	600
cacaacctga	ggaacagctg	ctgggtacog	atttgtattc	ctgtatatatac	tcgacttggg	660
tagccaggat	tttctcggtt	gtcttgaata	tattttacttc	taaaaagact	gttagtcttg	720
ttgggaggtg	agcagaggaa	tgtttacttg	ctggagaaca	tagtaacaaa	catcacttct	780

cacccctcat aaacagaaag cccattgtct gtgccgtgaa aagaggacta ccttgtcagg 840
 aagattttac ttacctctac aagaaaaagg actcgtatta aaatcgacct gtggttacca 900
 cccttagaaa tgctgttttc agatattcct attcctagtc ttagaggcag ccatggaata 960
 tatagtctcc tctgacctg gggtcaggaa gaaagggaaa tctattcata aaaggaattt 1020

<210> 382

<211> 668

<212> DNA

<213> Homo sapiens

<300>

<308> Contig35251_RC

<400> 382

ctatacagta tttgtcagct gtttttattc ttcagtttat catcttatat ttggctctca 60
 atttcaccca taaattattg atggcttatt atacgccaaag taccctgcgc tccgcttttt 120
 gtggatatccc tcttcggagc gagaagtcac ttagaatttt accccatctt cccacttgg 180
 gaatcggaag ggttcatagc taagttttgt tatttataag cagcccggt cttctctatc 240
 cctctctcca ccaaccccat ttcaagggtg gtccctcca gcggacgaca gtgatgaatt 300
 cctcttgggg acagccctcc cgcaactaac ggtggctcca tcgacgctct ggtctccttg 360
 ctgctcgtcg ggaattcggg tgggacgttt tccaaagtgt ggattccacg gcacattcct 420
 ccctcggaag ggcacaaccg gctcggcgtg tcagctcagc tggactcacg ccggtgcggg 480
 gccgcagctc cggttctcgc tgggcggctg cggagggtcg ggcgattctc tcagcgtcgc 540
 gaggcgaccc gggcagcagc aaacttctcc gcttccctc ttcgggacgc ggcacctctt 600
 ctgcccttga gcaactccga ggaacttccc ccggcctcgg tctcgtctcg aaataaccga 660
 gccagctc 668

<210> 383

<211> 430

<212> DNA

<213> Homo sapiens

<300>

<308> Contig25991

<400> 383

agaggagtcg gcgtttgaag aggtggaact cctagggctt ttttgagagt gctgatttag 60
 aagaatacaa atcatggctg aaaatagtgt attaacatcc actactggga ggactagctt 120
 ggcagactct tccatttttg attctaaagt tactgagatt tccaaggaaa acttacttat 180
 tggatctact tcatatgtag aagaagagat gcctcagatt gaaacaagag tgatattggg 240
 tcaagaagct ggaaaacaag aagaacttat aaaagcctta aaggacatta aagtgggctt 300
 tgtaaagatg gagtcagtgg aagaatttga aggtttggat tctccggaat ttgaaaatgt 360
 atttgtagtc acggactttc aggattctgt ctttaatgac ctctacaagg ctgattgtag 420
 agttattgga 430

<210> 384

<211> 2212

<212> DNA

<213> Homo Sapiens

<300>

<308> NM_003875

<400> 384

tgccggtctc tcctcgacca ggccctcttc tcaacctcag cccgcggcgc cgacccttcc 60
 ggcaccctcc cgccccgtct cgtactgtcg ccgtcacgc cgcggtccg gccctggccc 120
 cgatggctct gtgcaacgga gactccaagc tggagaatgc tggaggagac ctttaaggatg 180
 gccaccacca ctatgaagga gctgttgtca ttctggatgc tgggtgctcag tacgggaaag 240
 tcatagaccg aagagtgagg gaactgttcg tgcagtctga aattttcccc ttggaaacac 300
 cagcatttgc tataaaggaa caaggattcc gtgctattat catctctgga ggaccttaatt 360

```

ctgtgtatgc tgaagatgct cctgtggtttg atccagcaat attcactatt ggcaagcctg 420
ttcttggaat ttgctatggt atgcagatga tgaataaggt atttgagggt actgtgcaca 480
aaaaaagtgt cagagaagat ggagttttca acattagtgt ggataatata tgttcattat 540
tcaggggcct tcagaaggaa gaagttgttt tgcttacaca tggagatagt gtagacaaag 600
tagctgatgg attcaagggt gtggcacgtt ctggaaacat agtagcaggc atagcaaagt 660
aatctaaaaa gttatatgga gcacagttcc accctgaagt tggccttaca gaaaatggaa 720
aagtaatact gaagaatttc ctttatgata tagctggatg cagtggaaacc ttcaccgtgc 780
agaacagaga acttgagtgt attcgagaga tcaaagagag agtaggcacg tcaaaagtgt 840
tggttttact cagtgtgga gtagactcaa cagtttgtag agctttgcta aatcgtgctt 900
tgaaccaaga acaagtcat gctgtgcaca ttgataatgg ctttatgaga aaacgagaaa 960
gccagtcctgt tgaagaggcc ctcaaaaagc ttggaattca ggtcaaagtg ataaatgctg 1020
ctcattcttt ctacaatgga acaacaacc taccaatctc agatgaagat agaaccacac 1080
ggaaaaaat tagcaaaacg ttaaatatga ccacaagtcc tgaagagaaa agaaaaatca 1140
ttggggatag ttttgtaaag attgccaatg aagtaattgg agaaatgaac ttgaaaccag 1200
aggaggtttt ccttgcccaa ggtactttac ggctgatctt aattgaaagt gcatcccttg 1260
ttgcaagtgg caaagctgaa ctcatacaaa cccatcaca tgacacagag ctcacagaa 1320
agttgagaga ggagggaaaa gtaatagaac ctctgaaaga ttttcataaa gatgaagtga 1380
gaattttggg cagagaactt ggacttccag aagagttagt ttccaggcat ccatttccag 1440
gtcctggcct ggcaatcaga gtaatatgtg ctgaagaacc ttatatattgt aaggactttc 1500
ctgaaaccaa caatattttg aaaatagtag ctgatttttc tgcaagtgtt aaaagccac 1560
ataccctatt acagagagtc aaagcctgca caacagaaga ggatcaggag aagctgatgc 1620
aaattaccag tctgcattca ctgaatgcct tcttgctgcc aattaaaact gtaggtgtgc 1680
agggtgactg tcgttcctac agttacgtgt gtggaatctc cagtaaagat gaacctgact 1740
gggaatcact tatttttctg gctaggctta tacctcgcat gtgtcacaa gttaacagag 1800
ttgtttatat atttggccca ccagttaaag aacctcctac agatgttact cccactttct 1860
tgacaacagg ggtgctcagt actttacgcc aagctgattt tgaggcccat aacattctca 1920
gggagtctgg gtatgctggg aaaatcagcc agatgccggt gattttgaca ccattacatt 1980
ttgatcgga cccacttcaa aagcagcctt catgccagag atctgtgggt attcgaacct 2040
ttattactag tgacttcatt actggtatac ctgcaacacc tggcaatgag atccctgtag 2100
agggtgtatt aaagatggtc actgagatta agaagattcc tggattttct cgaattatgt 2160
atgacttaac atcaaaagcc ccaggaacta ctgagtggga gtaataaact tc 2212

```

<210> 385

<211> 2150

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006101

<400> 385

```

ctcgagccag gaaggcccg ctgtcctgtc tagcagatac ttgcacgggt tacagaaatt 60
cggtcctcgg gtcgtgtcag gaaactggaa aaaaggctcat aagcatgaag cgcagttcag 120
tttccagcgg tggtgctggc cgcctctcca tgcaggagtt aagatcccag gatgtaaata 180
aacaaggcct ctatacccct caaaccaaag agaaaccaac ctttgaaaag ttgagtataa 240
acaaaccgac atctgaaaga aaagtctcgc tattttggca aagaactagt ggacatggat 300
cccggaaatag tcaacttggg atattttcca gttctgagaa aatcaaggac ccgagaccac 360
ttaatgacaa agcattcatt cagcagtgtg ttcgacaact ctgtgagttt cttacagaaa 420
atggttatgc acataatgtg tccatgaaat ctctacaagc tccctctgtt aaagacttcc 480
tgaagatctt cacatttctt tatggcttcc tgtgcccctc atacgaactt cctgacacaa 540
agtttgaaga agaggttcca agaactctta aagaccttgg gtatcctttt gcactatcca 600
aaagctccat gtacacagtg ggggctcctc atacatggcc tcacattgtg gcagccttag 660
tttggtctaat agactgcac aagatacata ctgccatgaa agaaagctca cctttatttg 720
atgatgggca gccttgggga gaagaaactg aagatggaat tatgcataat aagttgtttt 780
tggaactacac cataaaatgc tatgagagtt ttatgagtgg tgccgacagc tttgatgaga 840
tgaatgcaga gctgcagtca aaactgaagg atttatttaa tgtggatgct tttaagctgg 900
aatcattaga agcaaaaaac agagcattga atgaacagat tgcaagattg gaacaagaaa 960
gagaaaaaga accgaatcgt ctagagtcgt tgagaaaact gaaggcttcc ttacaaggag 1020
atgttcaaaa gtatcaggca tacatgagca atttggagtc tcattcagcc attcttgacc 1080
agaaattaa tggtctcaat gaggaattg ctagagtaga actagaatgt gaaacaataa 1140

```

```

aacaggagaa cactcgacta cagaatatca ttgacaacca gaagtactca gttgcagaca 1200
ttgagcgaat aaatcatgaa agaaatgaat tgcagcagac tattaataaa ttaaccaagg 1260
acctggaagc tgaacaacag aagttgtgga atgaggagtt aaaatatgcc agaggcaaaag 1320
aagcgattga aacacaatta gcagagtatc acaaattggc tagaaaatta aaacttattc 1380
ctaaagggtgc tgagaattcc aaagggttatg actttgaaat taagtttaat cccgaggctg 1440
gtgccaaactg ccttgtcaaa tacagggtctc aagtttatgt acctcttaag gaactcctga 1500
atgaaactga agaagaaatt aataaagccc taaataaaaa aatgggtttg gaggatactt 1560
tagaacaatt gaatgcaatg ataacagaaa gcaagagaag tgtgagaact ctgaaagaag 1620
aagttcaaaa gctggatgat ctttaccac aaaaaattaa ggaagcagag gaagaggatg 1680
aaaaatgtgc cagtgaactt gagtccttg agaaacacaa gcacctgcta gaaagtactg 1740
ttaaccaggg gctcagtga gctatgaatg aattagatgc tgttcagcgg gaataccaac 1800
tagttgtgca aaccacgact gaagaaagac gaaaagtggg aaataacttg caacgtctgt 1860
tagagatggt tgctacacat gttgggtctg tagagaaaca tcttgaggag cagattgcta 1920
aagttgatag agaatatgaa gaatgcattg cagaagatct ctcggaaaat attaaagaga 1980
ttagagataa gtatgagaag aaagctactc taattaagtc ttctgaagaa tgaagataaa 2040
atgttgatca tgtatatata tccatagtga ataaaattgt ctcagtaaaa aaaaaaaaaa 2100
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2150

```

<210> 386
 <211> 2830
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_003882

```

<400> 386
cccacgcgtc cgctgggccc agctcccccg agaggtggtc ggatcctctg ggctgctcgg 60
tcgatgcctg tgccactgac gtccaggcat gaggtggctc ctgccctgga cgctggcagc 120
agtgcacagca gcagccgccca gcaccgtcct ggccacggcc ctctctccag cccctacgac 180
catggacttt actccagctc cactggagga cacctcctca cgcccccaat tctgcaagtg 240
gccatgtgag tgcccgcctc ccccaccccg ctgcccgctg ggggtcagcc tcatcacaga 300
tggtgtgag tgctgtaaga tgtgcgctca gcagcttggg gacaactgca cggaggctgc 360
catctgtgac ccccaccggg gcctctactg tgactacagc ggggaccgcc cgaggtacgc 420
aataggagtg tgtgcacagg tggctcgggtg gggtgcgtc ctggatgggg tgcgtacaa 480
caacggccag tccttccagc ctaactgcaa gtacaactgc acgtgcacg acggcgcggt 540
gggtgcaca ccactgtgcc tccgagtgcg ccccccgcgt ctctggtgcc cccaccgcg 600
gcgcgtgagc atacctggcc actgctgtga cactgtggta tgtgaggacg acgccaagag 660
gccacgcaag accgcacccc gtgacacagg agccttcgat gctgtgggtg aggtggaggc 720
atggcacagg aactgcatag cctacacaag cccctggagc ccttgctcca ccagctgcg 780
cctgggggtc tccactcgga tctccaatgt taacgcccag tgctggcctg agcaagagag 840
ccgcctctgc aacttgccgc catgcgatgt ggacatccat aactcatta aggcaggga 900
gaagtgtctg gctgtgtacc agccagaggc atccatgaac ttcacacttg cgggctgcat 960
cagcacacgc tcctatcaac ccaagtactg tggagtgtg atggacaata ggtgctgcat 1020
cccctacaag tctaagacta tcgacgtgct cttccagtgt cctgatgggc ttggcttctc 1080
ccgccaggtc ctatggatta atgcctgctt ctgtaacctg agctgtagga atcccaatga 1140
catctttgct gacttggaa cctaccctga cttctcagaa attgccaact aggcaggcac 1200
aaatcttggg tcttggggac taaccaatg cctgtgaagc agtcagccct tatggccaat 1260
aacttttcac caatgagcct tagttaccct gatctggacc cttggcctcc atttctgtct 1320
ctaaccattc aaatgacgcc tgatggtgct gctcaggccc atgctatgag ttttctcctt 1380
gatatcatto agcatctact ctaaagaaaa atgcctgtct ctagctgttc tggactacac 1440
ccaagcctga tccagccttt ccaagtcact agaagtcctg ctggatcttg cctaaatccc 1500
aagaaatgga atcaggtaga cttttaatat cactaatttc ttctttagat gccaaaccac 1560
aagactcttt ggggtccattc agatgaatag atggaatttg gaacaataga ataacttatt 1620
atttgagacc tgccaagagg tactgtaatg ggtaattctg acgtcagcgc accaaaacta 1680
tcttgattcc aaatatgtat gcacctcaag gtcacaaac atttgccaa tgagtgaat 1740
agttgcttaa ttttgatttt taatggaaa gtttatccat taacctgggc attgttgagg 1800
ttaagtttct cttcaccctt acactgtgaa gggtagagat taggtttgtc ccagtcagaa 1860
ataaaatttg ataaacattc ctgttgatgg gaaaagcccc cagttaatac tccagagaca 1920
gggaaaggtc agccacattc agaaggacca attgactctc aactgaatc agctgctgac 1980
tggcagggct ttgggcaggtt ggccaggctc ttccttgaat cttctccctt gtcctgcttg 2040

```

```

ggttcatagg aattggtaag gcctctggac tggcctgtct ggccctgag agtgggtgcc 2100
tggaacactc ctctactctt acagagcctt gagagaccca gctgcagacc atgccagacc 2160
cactgaaatg accaagacag gtacaggtag ggggtgtgggt caaaccaaga agtgggtgcc 2220
cttggttagca gcctggggtg acctctagag ctggaggctg tgggactcca ggggcccccg 2280
tgttcaggac acatctattg cagagactca tttcacagcc tttcgttctg ctgaccaa 2340
ggccagtttt ctggtaggaa gatggagggt taccagttgt ttagaaacag aaatagactt 2400
aataaagggt taaagctgaa gaggttgaag ctaaaaggaa aaggttgttg ttaatgaata 2460
tcaggctatt atttattgta ttaggaaaat ataataatta ctgttagaat tcttttattt 2520
agggcctttt ctgtgccaga cattgctctc agtgctttgc atgtattagc tcaactgaatc 2580
ttcacgacaa tgttgagaag ttcccattat tatttctggt cttacaaatg tgaaacggaa 2640
gctcatagag gtgagaaaac tcaaccagag tcaccaggtt ggtgactggg aaagtttaga 2700
ttcagatcga aattggactg tctttataac ccatattttc cccctgtttt tagagcttcc 2760
aaatgtgtca gaataggaaa acattgcaat aaatggcttg attttttaaa aaaaaaaaaa 2820
aaaaaaaaa 2830

```

<210> 387

<211> 2785

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003607

<400> 387

```

tccgcaagcc tccgcctctg tgcgcgggac tggaggagcc tcgctgagcc cagggcgcgca 60
cgccgagagg agaggggaagg cgggggaggg ctggaaggga gaggaaggga gtgggttgga 120
gccgggctgc cgcagcctct agtctctca gccgcggaag caccctcct cctgcgcgc 180
ggccgcctcc ctctcgctg tggaaagatg ccttagccc aggggtgtga agaaggggga 240
gaagtagctg ccagagccgc ccgcgccgcc gccgcagctg ctgctcgttg tgtccttga 300
attcagagaa gcaccccccc gcctggggcg tcgggagcct ccgggcggcg gccgcggtgc 360
ggtgtcgggg agaccgggct ctctgccgcg gcggcgcgcc gcggctcgcc ccacgagcga 420
ccaccgacat ggagtgggct cgggcggcca agtagccgct tctccggagc cgggtgcagt 480
gccgcccgca gccgccttc caccgccgc cgcgcgcccg gtcaggccct aggggtgaagc 540
cgggaggaaa atgaagagtt ttcaccggga atccgttgaa aataggactg actgcaaagc 600
cttaaagaaa gaaggacctc gggaggagaa acgaaaagcc gcctccgggc aagacttggc 660
gtgctccgag ccgaggggct gcttcaggga cctcgcccc tccctttccc gctggagaaa 720
ttgccgctga tgcattatcc aagtgggtgt tgggaggatt tgcagcaaca tttttggttt 780
tccctcccc ttctatgcat tctgtttttt tctcccttt gatgcaaccg gattcgtttc 840
gaagtgaatt gctgatgcaa atcggacttt attcattaat gatgcaaccg gattcgtttc 900
aggattacgt tgcacgagtt gaattttgaa tgaaggagaa gagttttttt ttttttttaa 960
agaagtgttg actctctagt tcgttgtact ttaattatt atttatttta aatatacgac 1020
ttaattgtat tcttttaaaa atgcattaag tatataattt atggtaattt accctcaaaa 1080
tatatgtata tgggtgaaat tgaagacgcc ttcagttaag tgaggttact ggtgtgttgg 1140
atgtttaatt cagcaccagc attgcatgac agttgtttga ataacaagtg gtttattttt 1200
aaaaccatac cttttaaaat ttaggttcag ataatagtaa aagtcatcat aataatttaa 1260
aggaaaacca gcagaaatcg aagcaaacat gtctggagaa gtgcgtttga ggcagttgga 1320
gcagtttatt ttggaogggc ccgctcagac caatgggcag tacttcagtg tggagacatt 1380
actggatata ctcatctgcc tttatgatga atgcaataat tctccattga gaagagagaa 1440
gaacattctc gaatacctag aatgggctaa accatttact tctaaagtga aacaaatgcg 1500
attacataga gaagactttg aaatattaaa ggtgattggg cgaggagctt ttggggaggt 1560
tgctgtagta aaactaaaaa atgcagataa agtgtttgcc atgaaaatat tgaataaatg 1620
ggaaatgctg aaaagagctg agacagcatg ttttcgtgaa gaaagggatg tattagttaa 1680
tggagacaat aattggatta caaccttgca ctatgcttcc caggatgaca ataacttata 1740
cctgggttatg gaattattat ttgggtggga tttgcttact ctactcagca aatttgaaga 1800
tagattgcct gaagatatgg ctagatttta cttggctgag atgggtgatg caattgactc 1860
agttcatcag ctacattatg tacacagaga cattaaacct gacaatatac tgatggatat 1920
gaatggacat attcggttag cagattttgg ttcttgtctg aagctgatgg aagatggaac 1980
ggttcagtc tcaagtggctg taggaactcc agattatata tctcctgaaa tcttcaagc 2040
catggaagat ggaagggga gatatggacc tgaatgtgac tgggtgtctt tgggggtctg 2100
tatgtatgaa atgctttacg gagaacacc attttatgca gaatcgctgg tggagacata 2160
cggaaaaatc atgaaccaca aagagaggtt tcagtttcca gcccaagtga ctgatgtgtc 2220

```

```

tgaaaatgct aaggatctta ttcgaaggct catttgtagc agagaacatc gacttgggtca 2280
aaatggaata gaagacttta agaaacaccc atttttcagt ggaattgatt gggataatat 2340
tcggaactgt gaagcacctt atattccaga agtttagtagc ccaacagata catcgaattt 2400
tgatgtagat gatgattgtt taaaaaattc tgaaacgatg cccccaccaa cacatactgc 2460
attttctggc caccatctgc catttgttgg ttttacatat actagtagct gtgtactttc 2520
tgatcggagc tgtttaagag ttacggctgg tcccacctca ctggatcttg atgttaatgt 2580
tcagaggact ctagacaaca acttagcaac tgaagcttat gaaagaagaa ttaagcgcct 2640
tgagcaagaa aaacttgaa ctagtagaaa acttcaagag tcaacacaga ctgtccaagc 2700
tctgcagtat tcaactgttg atggtccact aacagcaagc aaagatttag aaataaaaaa 2760
cttaaaagaa gaaatttgaa aaaaa 2785

```

<210> 388
 <211> 1912
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AF073519

```

<400> 388
ccggcggctg ttgtcgggcc tccagcgggc ggggcggtg gcggagcaga gcggaggcgc 60
agccggggcg agggcccacg agggctcagc cttcccggtc agcggtggtg acggtatccc 120
agagtgccag agaaccgttg cttttccgag ttgctottct tccaggtccc gttggtgggc 180
ggcatggccc gtggaaatca acgagaactt gcccgccaga aaaacatgaa gaaaaccag 240
gaaattagca agggaaagag gaaagaggat agcttgactg cctctcagag aaagcagagt 300
tctggaggcc agaaatctga gagcaagatg tcagctgggc cacacctccc tctgaaggct 360
ccaagggaga atccttgctt tctcttcca gctgctgggt gctccaggta ttacttggct 420
tatggcagca taactcctat ctctgccttt gtctttgttg tcttcttttc tgtcttcttc 480
ccttcttttt atgaggactt ttgctgttgg atttaggttc cattctaacc taggatgata 540
tcatttgtaa atccttaatt tcatctacaa aaactgtttt cccaaatagg tcacattcac 600
gcatatcaga tggacagatg tatcattttg gggccaacca ttcaaccac tacaaggagt 660
tttttaaaaca aaaataggaa acttagatgt aacttagcac tttttttttt tttttttgag 720
atggagtctc actctgtcac cagactggag tgcagtggcg ccatctcagc tccatgcaac 780
ctctgcctcc tgggttcaag cagttctctt gcctcagcct cctgggtagc tgggattaca 840
ggcacgcgct gccacaccca ggtaatttat ttattttttt tttgagacag agtctcgcac 900
tgttgcccag gctggactgc agtggcgtga tctctgctca ctgcaacctc cgcctcccg 960
gttcaagcga ttctocagcc tcagcttcct gagtagatgg gattacaggc gcctgccacc 1020
acgcccagct aatttttttg tattcttagt agagatgggg tttcaccatg ttggccaggc 1080
tgggtctccat ctctgacct cgtgattcac ccgcctcggc ctcccaaagt gctgggatta 1140
caggcgtgag tcacagcccc cggccataat tttagcattt aaaaaataat agccatgttg 1200
ggcaggcgtt ggtggctcat gcctgtaatc tgagcacttt gggagaccaa ggcggttaga 1260
tcccttgtgc ccaggagtcc aagaccagcc tgggcaacat ggcgaaaccc catttctact 1320
aaaaatacaa aaattagctg gggcgagggg ataggccgag ttccgggtgt aagggggcca 1380
ttagggagag cagagcgagg cagctgatct tccggttggt gggccttgcc cggaagctgg 1440
acctcacgga gatgaaacgg aagatgcacg aggatatgat ctccatacag aactttctca 1500
tctacgtggc cctgctgcga gtcactccat ttactttaa gaaattggac agcatatgaa 1560
gattggacat cacatgtgaa tgcatgatat gaagagcctg gttacagttt ctactgttct 1620
ctgcaagtaa ataggccag aaaggtataa gagactcttt gaatggacat aaaaattctg 1680
cttgtaaga acaagttgag ctctggtaac tgatcttaat agctaaaata taaaaatatt 1740
tgggaagtct gaaatgaggt ctctggccc tgggtgtgccc ttaatgcctg tgacagttgg 1800
cctctgtgaa tattggtata attgtaaata atgtcaaaact ccattttcta gcaagtatta 1860
ataattaagg gaagtatgtc tgaaatggca aaaaaaaaaa aaaaaaaaaa aa 1912

```

<210> 389
 <211> 1266
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_000849

<400> 389

```

ctcgggaagcc cgtcaccatg tcgtgcgagt cgtctatggt tctcgggtac tgggatattc 60
gtgggctggc gcacgccatc cgctgctcc tggagttcac ggatacctct tatgaggaga 120
aacggtacac gtgcggggaa gctcctgact atgatcgaag ccaatggctg gatgtgaaat 180
tcaagctaga cctggacttt cctaactctgc cctacctcct ggatgggaag aacaagatca 240
cccagagcaa tgccatcttg cgctacatcg ctcgcaagca caacatgtgt ggtgagactg 300
aagaagaaaa gattcagatg gacatcatag agaaccaagt aatggatttc cgcacacaac 360
tgataaggct ctgttacagc tctgaccacg aaaaactgaa gcctcagtac ttggaagagc 420
tacctggaca actgaaacaa ttctccatgt ttctgtggaa attctcatgg tttgccgggg 480
aaaagctcac ctttgtggat tttctcacct atgatatctt ggatcagaac cgtatatattg 540
accccaagtg attggatgag ttoccaaacc tgaaggcttt catgtgcogt tttgaggctt 600
tggagaaaaa cgctgcctac ttacagtctg atcagttctg caagatgccc atcaacaaca 660
agatggccca gtggggcaac aagcctgtat gctgagcagg aggcagactt gcagagcttg 720
ttttgtttca tctgttccgt aaggggtcag cgctcttgct ttgctctttt caatgaatag 780
cacttatgtt actggtgtcc agctgagttt ctcttgggtg taaaggctaa aagggaataa 840
ggatatgttg agaatcatca agatatgaat tgaatcgctg cgatactgtg gcatttccct 900
actccccaac tgagttcaag ggctgtaggt tcatgcccac gccctgagag tgggtactag 960
aaaaaacgag attgcacagt tggagagagc aggtgtgtta aatggactgg agtcctctgtg 1020
aagactgggt gaggataaca caagtaaaac tgtggtagtg atggacttaa ccggagtctg 1080
gaaaccgtcc tgtgtacaca tgggagttta gtgtgataaa ggcagtattt cagactgggtg 1140
ggctagccaa tagagttggc aattgcttat tgaactcat taaaaataat agagcccac 1200
ttgacactat tcaactaaat taatctggaa ttaaggccc aacattaaac acaaagctgt 1260
attgat 1266

```

<210> 390

<211> 517

<212> DNA

<213> Homo sapiens

<300>

<308> Contig32185_RC

<400> 390

```

ttaattagag tttgtagtat aacttaact ttctgtaato tgcagtgatg aatctctatg 60
taaacattca gaaaaagagc gaatactgtg tcatgactta tgaactataa attttgtccc 120
tgattactag tcaagaacag tttatacat ttattaacaa ctactttcct ccacttaaca 180
acagctctta aataacaaaa taaattatca gcatttccaa ccacaaattt aagaattaaa 240
gtttaaaaaa ctggttagtt tcataaatgc aagcagagta ataaatttca tatatgttaa 300
ttgcagttac ttatgttctt tgagtgaagt aacacaaaga tttttccctt ccccttttgg 360
caagttcaca aaattcatga aaagtcttca cctttttctt ttttgcttaa tttctttttt 420
gcattctcagc tgaacacaac agacaatagc tggggttatt agcgagacat gataaaagaa 480
aaggttatga ccactcagac gatcatgtat acctaaa 517

```

<210> 391

<211> 1266

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016577

<400> 391

```

gcgcgcgcgc cctgcgcgcc gctccgcgcc cgctctcttc ccagcgcgcg gcctcggggc 60
taaagagcgc gccctcgcga ccgcagccag tgcgcagcgc agcagcccag cccagcctt 120
cctccgcctc cgctctcttc tctcctcct cctccgcgcg cggacacgca gccgcagccg 180
ggaccgggac gcagctgggg agtcagggac gcgcgcagcc agcccttccc cctccggctc 240
ccgcaccgcc ggccgcctcc cctcgccttc ctactctccc ctccctgctc ctccgctttt 300
tctcctcct cctctccogg ccccggtgc cagcaccatg tccgcagggg gagattttgg 360
gaatccactg agaaaattca agttggtgt cttgggggag cagagcgtcg ggaagacgtc 420
tctgattacg aggttcatgt acgacagctt cgacaacaca taccaggcaa ccattgggat 480
tgacttcttg tcaaaaacca tgtacttggg ggaccgcacg gtgcgactgc agctctggga 540

```

```

cacagctggt caggagaggt tccgcagcct gatccccagc tacatccggg actccacggt 600
ggctgtggtg gtgtacgaca tcacaaatct caactccttc caacagacct ctaagtggat 660
cgacgacgtc aggacagaga ggggcagtga tggtatcatc atgctggtgg gcaacaagac 720
ggacctggct gataagaggc agataacctat cgaggagggg gagcagcgcg ccaaagaact 780
gagcgtcatg ttcatcgaga ccagtgcgaa gactggctac aacgtgaagc agctttttcg 840
acgtgtggcg tcggtctctac ccggaatgga gaatgtccag gagaaaagca aagaagggat 900
gatcgacatc aagctggaca aaccccgagg gcccccggcc agcgaggggc gctgctcctg 960
ctaatacgaga gccgacctgt ggcttcccat gacactcctt gcttggttggt ttgcttccta 1020
ttgctagct tcctaagggg ggaggggaacc gagttatcaa gatgggagga tttttctttt 1080
ctctctgtct ttaggagtag ggtgggatgg ggaggaggc tgggcatcag ggatcacatc 1140
actcttaacg gctgttactt aaacaactat tttttggtt gggtgtaata tattgtactt 1200
tattaagatt gccaaaaact gttaaaattt aaaaaaatt taaatcatgt gtatacaatt 1260
ttttgc

```

<210> 392
 <211> 472
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig48328_RC

```

<400> 392
gtacttttat tgaaaaggta catttaaaaa aatacacaga cattttacca ttacaggtt 60
gcagatatag atgctctaaa agagtccact ctattttgtt gttctatgat aactcttgcc 120
cctgatatca caaacattcc agtcttggtg atatcggtt agaaaggggg gcatgggagc 180
atgacctgca atataattcag cgaacagaaa gacaaattgt tcaattataa atttttttat 240
cttctgtaca ttattgcatt agggagccac aaaattatgt agcatcatta caaatgaaaa 300
caggttaaaa atgaagaaga tacttatata gaaatacatg gattcattgt cttcttgagc 360
aatgcacaag aggtgcaaaa atgtgcaatt taggaagctc tttttctgtt tgtatacgtt 420
tgcttagcaa cacaaaccag tgaggaagct acaaaataag ttaaacaaaa at 472

```

<210> 393
 <211> 431
 <212> DNA
 <213> Homo sapiens

<300>
 <308> Contig46223_RC

```

<400> 393
cttgtccaag atcattttatt gttattatta tttttctctc ttaattggat tcagcgtttc 60
tggctcctca cacacatcct ctctggtggg ggattcaggt ttgtgcaaag cacaaggag 120
cttgtaacca cctgggcccc cgccagcag cacggccggc ttgatcgga gccctcttc 180
cctgtccggg ctgcctgccc ccagtccctg ctctgacccc ctgagtcagg cctcccagc 240
ctctgcctgc accctgagca tccccaggac ggctgggagg gccaggctct ccgtgatgca 300
gaagtgcctg agatggttca tgaaatgcc tcattgttcc ccttctgcta aggcccgcg 360
acgcctccct cctgctgtcc cacacagggc ccttgctggg tcacacacca aacagcctca 420
aggaaaaatt c
431

```

<210> 394
 <211> 1728
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_015984

```

<400> 394
cgggaagtag gcggggtgac gtgtggttga cgagctcggc ggcggggttg ctgagatctg 60
tggccggcgg cagctggtgc ggggggcagc tgagagcgag aggtggatcg ggcgggtgtg 120

```

```

tggccagggc catgacgggc aacgccgggg agtgggtgcct catggaaagc gaccccgggg 180
tcttcaccga gctcattaaa ggattcgggt gccgaggagc ccaagtagaa gaaatatgga 240
gttttagagcc tgagaatttt gaaaaattaa agccagttca tgggttaatt tttcttttca 300
agtggcagcc aggagaagaa ccagcaggct ctgtggttca ggactcccga cttgacacga 360
tattttttgc taagcaggta attaataatg cttgtgctac tcaagccata gtgagtgtgt 420
tactgaactg taccaccag gatgtccatt taggcgagac attatcagag tttaaagaat 480
tttcacaaag ttttgatgca gctatgaaag gcttggcact gagcaattca gatgtgattc 540
gacaagtaca caacagtttc gccagacagc aaatgtttga atttgatacg aagacatcag 600
caaaagaaga agatgctttt cactttgtca gttatgttcc tgtaaatggg agactgtatg 660
aattagatgg attaagagaa ggaccgattg atttaggtgc atgcaatcaa gatgattgga 720
tcagtcagtc aaggcctgtc atagaaaaaa ggatacaaaa gtacagtga ggtgaaattc 780
gatttaattt aatggccatt gtgtctgaca gaaaaatgat atatgagcag aagatagcag 840
agttacaaag acaacttgca gaggaggaac ccatggatac agatcaaggt aatagtatgt 900
taagtgtctat tcagtcagaa gttgccaaaa atcagatgct tattgaagaa gaagtacaga 960
aattaaaaag atacaagatt gagaatatca gaaggaagca taattatctg cctttcatta 1020
tggaattgtt aaagacttta gcagaacacc agcagttaat accactagta gaaaaggcaa 1080
aagaaaaaca gaacgcaaaag aaagctcagg aaaccaaagc aagatgtttt cagatatgta 1140
cacatttctg cttctgcaca tattttcatg gaaaccatta tgtataaaga acttagagca 1200
acatcctaatt tggctcagtg cacgtttggc aatagtgcc aacctgtctt tctttaatgc 1260
atggattcat aaacttcttc cctacctgca tcatgtgcat gtagtgcata ttaaatgaaa 1320
gtgatattaa gaatgctttc ccaaattcca ttatttgaca ttgagtctga caactgttag 1380
ttttctgggt gtccaactac catatgaagc tagaaaatgc acaaacgata ttccttatct 1440
gtaattttaa tacttaaaat ttgcaattgt cagatcttga ttaactgggt tgtcttattt 1500
cttctcatca ttaacggaaa aaaaatcagt atttctatct ttgatatcta agtgttttga 1560
ggatttttaa actgaatttt atctgctata ccagttattt gagaaagtat gattttaatg 1620
taaatcattt aaaaaggaca aaagtataat ttccagtgat tttcactgct gtcagtagaa 1680
aagtaataaa catctcaatt ttaaaaaaaa aaaaaaaaaa aaaaaaaa 1728

```

<210> 395

<211> 1348

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006117

<400> 395

```

cagggtggcgt acttggttg gagactggcg cggcggttcgt gtccgagttc tctgcaggtc 60
actagtttcc cggtagttca gctgcacatg aatagaacag caatgagagc cagtcagaag 120
gacttttgaaa attcaatgaa tcaagtgaag ctcttgaaaa aggatccagg aaacgaagtg 180
aagctaaaaac tctacgcgct atataagcag gccactgaag gaccttgtaa catgccaaa 240
ccagggtgat ttgacttgat caacaaggcc aaatgggacg catggaatgc ccttggcagc 300
ctgcccaagg aagctgccag gcagaactat gtggatttgg tgtccagttt gattccttca 360
ttggaatcct ctagtcagggt ggagcctgga acagacagga aatcaactgg gtttgaact 420
ctggtggtga cctccgaaga tggcatcaca aagatcatgt tcaaccggcc caaaaagaaa 480
aatgccataa acactgagat gtatcatgaa attatgcgtg cacttaaagc tgccagcaag 540
gatgactcaa tcatcactgt tttaacagga aatggtgact attacagtag tgggaatgat 600
ctgactaact tcactgatat tccccctggt ggagtagagg agaaagctaa aaataatgcc 660
gttttactga gggaatttgt gggctgtttt atagattttc ctaagcctct gattgcagtg 720
gtcaatggtc cagctgtggg catctcogtc accctccttg ggctattcga tgccgtgtat 780
gcatctgaca gggcaacatt tcatacacca tttagtcacc taggccaaaag tccggaagga 840
tgctcctctt acacttttcc gaagataatg agcccagcca aggcaacaga gatgcttatt 900
tttgaaaaga agttaacagc gggagaggca tgtgctcaag gacttggtac tgaagttttc 960
cctgatagca cttttcagaa agaagtctgg accaggctga aggcatttgc aaagcttccc 1020
ccaaatgcct tgagaatttc aaaagaggta atcaggaaaa gagagagaga aaaactacac 1080
gctgttaatg ctgaagaatg caatgtcctt cagggaagat ggctatcaga tgaatgcaca 1140
aatgctgtgg tgaacttctt atccagaaaa tcaaaactgt gatgaccact acagcagagt 1200
aaagcatgtc caaggaagga tgtgctgtta cctctgattt ccagtactgg aactaaataa 1260
gcttcattgt gccttttgta gtgctagaat atcaattaca atgatgatat ttcactacag 1320
ctctgatgaa taaaaagttt tgtaaaac 1348

```

<210> 396
 <211> 3106
 <212> DNA
 <213> Homo sapiens

<300>
 <308> AK000745

<400> 396
 cggccgcggc ccccgacgac ctggccttgc tgaagaatct ccggagcgag gaacagaaga 60
 agaagaaccg gaagaaactg tccgagaagc ccaaaccaaa tgggcggact gttgaagtgg 120
 ctgaggggtga agctgttcga acacctcaaa gtgtaacagc aaagcagcca ccagagattg 180
 acaagaaaaa tgaaaagtca aagaaaaata agaagaaatc aaagtcagat gctaaagcag 240
 tgcaaaacag ttcacgccat gatggaaagg aagttgatga aggagcctgg gaaactaaaa 300
 ttagtcacag agagaaacga cagcagcgta aacgtgataa ggtgctgact gattctgggt 360
 cattggattc aactatccct gggatagaaa ataccatcac agttaccacc gagcaactta 420
 caaccgcatac atttctctgtt gggtccaaga agaataaagg tgattctcat ctaaagtgtc 480
 aagttagcaa ctttaaatct ggaaaaggag attctacact tcaggtttct tcaggattga 540
 atgaaaacct cactgtcaat ggaggaggct ggaatgaaaa gtctgtaaaa ctctcctcac 600
 agatcagtgc aggtgaggag aagtggaaact ccgtttcacc tgcttctgca ggaaagagga 660
 aaactgagcc atctgcctgg agtcaagaca ctggagatgc taatacaaat ggaaaagact 720

 ggggaaggag ttggagtgac cgttcaatat tttctggcat tgggtctact gctgagccag 780
 tttctcagtc taccacttct gattatcagt gggatgttag ccgtaatcaa ccctatatcg 840
 atgatgaatg gtctgggtta aatgggtctgt cttctgctga tcccaactct gattggaatg 900
 caccagcaga agagtggggc aattgggttag acgaagaaag agcttcactt ctaaagtccc 960
 aggaaccaat tcctgatgat caaaaggctc cagatgatga taaagaaaag ggagagggag 1020
 ctcttccaac tgggaaatcc aaaaagaaaa aaaagaaaaa gaagaagcaa ggtgaagata 1080
 actctactgc acaggacaca gaagaattag aaaaagagat tagagaagac cttccagtga 1140
 atacctctaa aaccgcgtcca aaacaggaaa aagctttttc cttgaagacc ataagcacta 1200
 gtgateccagc cgaagtactc gtcaaaaaata gccagcctat caagactctt ccacctgcta 1260
 cttctaccga gccatctgta atcttatcaa aaagtgttcc tgacaagagc tcttcccaag 1320
 tgccgcgaat actacaagag acagataaat ccaagtcaaa taccagcaa aatagtgtgc 1380
 ctccctcaca gaccaagtct gaaactagct gggaaatctc caaacaataa aaaaagaaga 1440
 aaaaagccag acgagaaacg tgaaatTTTT tttcctgaat tggacatgtg tttgcaaaaa 1500
 cttgtcttga agattatgct gtttatgcaa taatttgtga acatgtacag agttttatat 1560
 aaattttaac caatttttaa aacaaaactg cggacaccac cataaaaatg gaatcaaaag 1620
 aaagttaatt tatgaaatta agaggtcagc agaataact cagtgtatga agacacttgg 1680
 gaaagtcttt ttaatagaac aagaacgata ttaatttaag aatattatcc tggtttaaca 1740
 acagtgcctt gtttacaaca gattgtgccc tatctcatct gcagccgagg aataaaggat 1800
 tctgtataga aagaggggtg cctacagatt agtaagcaat tccttggatc ttatgcacag 1860
 aacttgtacc atttgaatct gttttatgct taaatcaag tgctttgatc aaatgcataa 1920
 cctgccatat ctttaccatat ttgttggtag caatttgtat taaagaaatc acaagtgcag 1980
 ataaaaagtc atttatcatt tgtttaacta aactgtcatg gtttagttta caatttttaa 2040
 aaagtcttta aaatactgaa aatgcagttg acacttgtgt atggcttatg aagttatttt 2100
 tgatagtctt acattacttg aattgttcaa agtacagtat attttaaatt aagaaaagtg 2160
 aactatatgt atttgtttta tacattttaag gcttagactc ataaataatg ctattgttta 2220
 tgatttgaaa actttcaggc aaaatccaat ttacattttt cccttcctta gcaattactt 2280
 ttttccagct tcaactcttc ttagttacta atactttgtt gactttaaaa atgaaatcat 2340
 tcacaaactt ttggtatatg atggagaatg aaaaactaga gtcagacagc ttttaattgac 2400
 attgtcaaca cctccagtta tcaggaatac atttttttac tgccttaacc tgtagtgcgt 2460
 agaatatgca tcaatttctt gaaggagatt catgttttta taagaatttt catgtaatta 2520
 ttgcaattgt ggtcaaaata ggaacgtttc ctgcttgaaa ttatattgat ttaaatgatg 2580
 tgtgagatgt ttcaccattt tcaggcactg ttgaattcta ttgtaataaa ctggcagga 2640
 tctttgtaac tataaatagt gcatgctcag ccattgtcac tgtaaatagc ctttaccaaa 2700
 cgtgtttgac aaggaccata attaacatca cttagtgaat tgtgataaag aaaaaaagc 2760
 catgatttat tcgatgtgat tggcttgttt ttatgtggcg ccaagaacga acctgtttta 2820
 cagctgtaac caatggtact gatctatcca tccaatgttg tcattatatt tgactgtgg 2880
 tcaacagtat tgcgttgtca gactaggaaa gctaaacgaa caaatgggtt ttagttttgc 2940
 tgaagactgg ccttattaat ggacagcttt cctaacaaga gattattaac ttttatcagg 3000
 tgttaacatc tgtttcagga acatggcagc atgtttacat gtcagaagtt ttgtttaatt 3060

ctatgggtatt tctaaattga cttgttttaa taaattcagc aaatgg

3106

<210> 397

<211> 556

<212> DNA

<213> Homo sapiens

<300>

<308> Contig40831_RC

<400> 397

```

cttgttttat catgtaattt tcagaaaata tgtgatccat aaagattaaa agaaagtgtg 60
attaagtctg gcagcttttag tattaacttg aaataaaaata tggcaagctt tccacgtcct 120
cctttatttc cacaatccat atgtacgagc tagattccag tcagaacttc cacaataact 180
tcaactcttg gtagcagcgg ttataaatta cgcctttgct aatttgcgtt gttcccaacc 240
aggagaaaaca ttaccacaaa aaaagtcagt ttcatcctgc agtgttcccg cagcaaccat 300
attaagctgg aagaataaag ctcccttgta gtagtagcgt ctattgaatt acaatgtaaa 360
tgttgatgcc tccoctggaa acgtcaaattg acaacagggt tacagaacga attcagtggg 420
tgtttgagcgc tctggaatcc caggctgcag agttgagatc gaagagggtgc gtctggccat 480
cttagcgttc attcaaccaa atctcctcac tccatcaaaa ccacgccact cccacactgc 540
ccagctcgta aaagga

```

556

<210> 398

<211> 893

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003239

<400> 398

```

gagtttttagc aggaggggat atgcatgtac acatatgcat atacatacag gtttatttaa 60
gtgattggat ctacactttt ggaggctatt tgaagtatct tatcaaaaata aaacgcattt 120
taagcaagag gttaaaaaaag cttttcagac ccgaaggcca ttaatacag taaatottac 180
ttgggtagtt tagcaaacat tttttaaaat ccacatccaa cagattgggtt agtattagat 240
acacagattt ggtttcaaag ctgaaatacc cctggtggaa taaatcttct catgaaattc 300
agttatatta tgccacaaaa cagggacact tatctattga caacttaaat attaaactcag 360
taggctacct ggtgatataa ataggaaaaa aacagtgggtc atggattagg agtgagaaga 420
ctatgacacc agcctacctt ggactgggac cgcaaaggat gttaaaatct atcctagaat 480
ttaaaacagg aaaaatgtca aatattccac tccccattcg gcccaaacca tcccccaagg 540
aatacacccc cccccacccc ccacatgagg tttaactaacc agctgctaca gatgttacta 600
tgctactaat taacaatgat cacagaaaat gtgtcatacc agagtttttg ttatcataca 660
caatgcatta gtggctgcag gaatcttctc caaaagcta ccgtcagaag gcaaactgac 720
aaatgtccat gcogctggtc caagaaacca cagctaagag ccatgaggca ttcatTTgta 780
cattgagctc tctctggcat gctccccact ctagttagag gggctggggt atggagagac 840
cgtcccagcc gggcgctgca gcctcctact cctgggcctc ttgcgggtca act 893

```

<210> 399

<211> 2470

<212> DNA

<213> Homo sapiens

<300>

<308> NM_014791

<400> 399

```

ttggcgggcg gaagcggcca caaccggcg atcgaaaaga ttcttaggaa cgccgtacca 60
gccgcgtctc tcaggacagc aggccctgt ccttctgtcg ggcgcgctc agccgtgcc 120
tcgccctc aggttctttt tctaattcca aataaacttg caagaggact atgaaagatt 180
atgatgaact tctcaaatat tatgaattac atgaaactat tgggacagggt ggctttgcaa 240
aggtcaaact tgcttgccat atccttactg gagagatggt agctataaaa atcatggata 300

```

```

aaaacacact agggagtgat ttgccccgga tcaaaacgga gattgaggcc ttgaagaacc 360
tgagacatca gcatatatgt caactctacc atgtgctaga gacagccaac aaaatattca 420
tggttcttga gtactgccct ggaggagagc tgtttgacta tataatttcc caggatcgcc 480
tgtcagaaga ggagaccggy gttgtcttcc gtcagatagt atctgctgtt gcttatgtgc 540
acagccaggg ctatgctcac agggacctca agccagaaaa tttgctgttt gatgaatata 600
ataaattaaa gctgattgac tttggtctct gtgcaaaacc caagggtaac aaggattacc 660
atctacagac atgctgtggg agtctggcct atgcagcacc tgagttaata caaggcaaat 720
catatcttgg atcagaggca gatgtttgga gcatgggcat actgttatat gttcttatgt 780
gtggatttct accatttcat gatgataatg taatggcttt atacaagaag attatgagag 840
gaaaatatga tgttcccaag tggctctctc ccagtagcat tctgcttctt caacaaatgc 900
tgcaaggtaga cccaaagaaa cggatttcta tgaaaaatct attgaaccat ccttggtgca 960
tgcaagatta caactatcct gttgagtgcc aaagcaagaa tccttttatt cacctcgatg 1020

```

```

atgattgcgt aacagaactt tctgtacatc acagaaacaa caggcaaaca atggaggatt 1080
taatttcaact gtggcagtat gatcacctca cggctaccta tcttctgctt ctagccaaga 1140
aggctcgggg aaaaccagtt cgtttaaggc tttcttcttt ctctgtgga caagccagt 1200
ctacccatt cacagacatc aagtcaaata attggagtct ggaagatgtg accgcaagt 1260
ataaaaatta tgtggcggga ttaatagact atgattggtg tgaagatgat ttatcaacag 1320
gtgctgctac tcccgaaca tcacagttta ccaagtactg gacagaatca aatgggtggtg 1380
aatctaaatc attaaactcca gccttatgca gaacacctgc aaataaatta aagaacaaag 1440
aaaatgtata tactcctaag tctgctgtaa agaatagaaga gtactttatg tttcctgagc 1500
caaagactcc agttaataag aaccagcata agagagaaat actcactacg ccaaactcgtt 1560
acactacacc ctcaaaagct agaaaccagt gcctgaaaga aactccaatt aaaataccag 1620
taaattcaac aggaacagac aagttaatga caggtgtcat tagccctgag aggcggtgcc 1680
gctcagtgga attggatctc aaccaagcac atatggagga gactccaaaa agaaaggag 1740
ccaaagtgtt tgggagcctt gaaagggggt tggataaggt tatcactgtg ctcaccagga 1800
gcaaaaggaa ggggtctgcc agagacgggc ccagaagact aaagcttcac tataatgtga 1860
ctacaactag attagtgaat ccagatcaac tgttgaatga aataatgtct attcttccaa 1920
agaagcatgt tgactttgta caaaagggtt atacactgaa gtgtcaaaca cagtcagatt 1980
ttgggaaagt gacaatgcaa tttgaattag aagtgtgcc gcttcaaaaa cccgatgtgg 2040
tggttatcag gaggcagcgg cttaaggggc atgcctgggt ttacaaaaga ttagtggag 2100
acatcctatc tagctgcaag gtataattga tggattcttc catcctgccg gatgagtgtg 2160
ggtgtgatac agcctacata aagactgtta tgatcgcttt gattttaaag ttcattggaa 2220
ctaccaactt gtttctaaag agctatctta agaccaatat ctctttgttt ttaaacaaaa 2280
gatattatct tgtgtatgaa tctaaatcaa gccatctgtt cattatgtta ctgtcttttt 2340
taatcatgtg gttttgtata ttaataattg ttgactttct tagattcact tccatatgtg 2400
aatgtaagct cttaactatg tctctttgta atgtgtaatt tctttctgaa ataaaaccat 2460
ttgtgaatat 2470

```

<210> 400
 <211> 2212
 <212> DNA
 <213> Homo

<300>
 <308> X05610

```

<400> 400
ggggaacgag gcccacctgg gagcccagga cttcaggggt tcccaggcat cacaccccct 60
tccaacatct ctggggcacc tggtgacaaa ggggcgccag ggatatttgg cctgaaaggt 120
tatcggggcc caccagggcc accaggttct gctgctcttc ctggaagcaa aggtgacaca 180
gggaacccag gagctccagg aaccocaggg accaaaggat gggccgggga ctccggggcc 240
cagggcaggg ctggtgtgtt tgggtctcca ggagaaaaag ggcccagggg tgaacaaggc 300
ttcatgggga acactggacc caccggggcg actggcgaca gaggcccaa gggaccaaag 360
ggagaccag gattccctgg tgcctccggg actgtgggag ccccgggat tgcaggaatc 420
ccccagaaga ttgccatcca accaggggaca gtgggtcccc aggggaggcg agggccccct 480
ggggcaccgg gggagatcgg gcccagggc ccccgcgag aaccaggttt tctgtgggct 540
ccagggaaag ctgggcccc aaggaagggt ggtgtgtctg ctgttcccg cttccgggga 600
gatgaaggac ccataggcca ccaggggccg attggccaag aaggtgcacc aggccttcca 660
gggagcccg gctgcccggg tatgccaggc cgcagcgtca gcatcggtca cctcctggtg 720
aagcacagcc agacggacca ggagcccatg tgcccgtgg gcatgaacaa actctggagt 780

```

ggatacagcc tgctgtactt cgagggccag gagaaggcgc acaaccagga cctggggctg 840
gcgggctcct gcctggcgcg gttcagcacc atgcccttcc tgtactgcaa ccctgggtgat 900
gtctgtact atgccagccg gaacgacaag tctactggc tctctaccac tgcgccgtg 960
cccatgatgc ccgtggccga ggaacgagatc aagccctaca tcagccgctg ttctgtgtgt 1020
gagggcccg ccacgccat cgcggtccac agtcaggatg tctccatccc aactgccc 1080
gctgggtggc ggagtttgtg gatcgatat tcttctctca tgcacacggc ggcgggagac 1140
gaaggcggtg gccaatcact ggtgtcaccg ggcagctgtc tagaggactt ccgcgccaca 1200
ccattcatcg aatgcaatgg aggcgcggc acctgccact actacgccc 1260
ttctggctga ccaccattcc cgagcagagc ttccagggtc cgccctcgc cgacacgctc 1320
aaggccggcc tcacccgcac acacatcagc cgctgccagg tgtgcatgaa gaacctgtga 1380
gcccggcgct gccaggaagg gccatttttg tcttattct taacttatta cctcaggtgc 1440
caacccaaaa ttggttttat tttttctta aaaaaaaaa aaagtctacc aaaggaattt 1500
gcatccagca gcagcactta gacctgccag ccactgtcac cgagcgggtg caagcactcg 1560
gggtccctgg aggcagccc tgcccacaga aagccaggag cagccctggc ccccatcagc 1620
cctgtacga cgcaccgct gaaggcacag ctaaccactt cgcacacacc catgtaacca 1680
ctgcactttc caatgccaca gacaactcac attgttcaac tcttctcgg ggtgggacag 1740
acgagacaac agcacacagg cagccagccg tggccagagg ctcgaggggc tcaggggctc 1800
aggcaccgct cccacacga gggcccgctg ggtggcctgg ccctgcttc tacgccaatg 1860
ttatgccagc tccatgttct cccaaatacc gttgatgtga attatttta aggcaaaact 1920
gtgctcttta ttttaaaaa cactgataat cacactgcgg taggtcattc ttttgccaca 1980
tccctataga cactgggtt tggcaaaact caggcagaag tggagacctt tctagacatc 2040
attgtcagcc ttgtacttg aaggtacacc ccatagggc ggaggtgctg tccccactgc 2100
cccacctgt cctgagatt taaccctcc actgctggg gtgagctgta ctcttctgac 2160
tgccccctcc tgtgtaacga ctacaaaata aaacttggt ctgaatat tt 2212

<210> 401

<211> 4221

<212> DNA

<213> Homo sapiens

<300>

<308> NM_016448

<400> 401

cgataacgat ttgtgttgtg agaggcgcaa gctgcgattt ctgctgaact tggaggcatt 60
tctacgactt ttctctcagc tgaggctttt cctccgaccc tgatgctctt caattcgggtg 120
ctccgccagc ccagcttgg cgctcctgaga aatggatggt ctccacaata ccctcttcaa 180
tccctctga ctggttatca gtgcagtggt aatgatgaac acacttctta tggagaaaca 240
ggagtcccag ttctctcttt tggatgtacc ttctctctctg ctcccaatat ggaacatgta 300
ctagcagttg ccaatgaaga aggttttggc cgattgtata acacagaatc acaaaagtctc 360
agaaagaagt gcttcaaaaga atggatggct cactggaatg ccgtctttga cctggcctgg 420
gttctgtgtg aacttaaaact tgttacagca gcaggtgatc aaacagccaa attttgggac 480
gtaaaagctg gtgagctgat tggaaatgc aaaggtcatc aatgcagcct caagtcaagt 540
gccttttcta agtttgagaa agctgtatct tgtacgggtg gaagagatgg caacattatg 600
gtctgggata ccaggtgcaa caaaaaagat gggttttata ggcaagtga tcaaatcagt 660
ggagctcaca atacctcaga caagcaaac ccttcaaac ccaagaagaa acagaattca 720
aaaggacttg ctcttctgt ggatttocag caaagtgtta ctgtggtcct ctttcaagac 780
gagaatacct tagtctcagc aggagctgtg gatgggataa tcaaagtatg ggatttacgt 840
aagaattata ctgcttatcg acaagaaccc atagcatcca agtcttctct gtaccaggt 900
agcagcactc gaaaacttgg atattcaagt ctgatttttg attccactgg ctctacttta 960
tttgctaatt gcacagaaga taacatctac atgtttaata tgactgggtt gaagacttct 1020
ccagtggcta ttttcaatgg acaccagaac tctacctttt atgtaaaatc cagccttagt 1080
ccagatgacc agtttttagt cagtggctca agtgaaga ctgcctacat atggaaggtc 1140
tccacacctt ggcaacctcc tactgtgctc ctgggtcatt ctcaagaggt cactctgtg 1200
tgctgggtgtc catctgactt cacaaagatt gctacctgtt ctgatgacaa tacactaaaa 1260
atctggcgct tgaatagagg cttagaggag aaaccaggag gtgataaact ttccacgggtg 1320
ggttgggctt ctcaagaaga aaaagagtca agacctggc tagtaacagt aacagtagc 1380
cagagtactc ctgccaaagc ccccagggtg aagtgaatc catccaattc tccccgtca 1440
tccgagctt gtgccccaaag ctgtgctgga gacctccctc tctcttcaa tactcctacg 1500
ttctctatta aaacctctcc tgccaaggcc cggtctccca tcaacagaag aggtctgtc 1560
tctccgtct ctccaagcc accttcatct ttcaagatgt cgattagaaa ctgggtgacc 1620

```

cgaacacctt cctcatcacc acccatcact ccacctgctt cggagaccaa gatcatgtct 1680
ccgagaaaag cccttattcc tgtgagccag aagtcattcc aagcagaggg ttgctctgag 1740
tctagaaaata gagtaaagag gaggctagac tcaagctgtc tggagagtgt gaaacaaaag 1800
tgtgtgaaga gttgtaactg tgtgactgag cttgatggcc aagttgaaaa tcttcatttg 1860
gatctgtgct gccttgctgg taaccaggaa gaccttagta aggactctct aggtcctacc 1920
aaatcaagca aaattgaagg agctgggtacc agtatctcag agcctccgtc tcctatcagt 1980
ccgtatgctt cagaaaagctg tggaaacgcta cctcttcctt tgagaccttg tggagaaggg 2040
tctgaaatgg taggcaaaga gaatagttcc ccagagaata aaaactggtt gttggccatg 2100
gcagccaaac ggaaggctga gaatccatct ccacgaagtc cgtcatccca gacacccaat 2160
tccaggagac agagcggaaa gacattgccca agcccgggtc ccatcacgcc cagctccatg 2220
aggaaaatct gcacatactt ccatagaaaag tcccaggagg acttctgtgg tcctgaacac 2280
tcaacagaat tatagattct aatctgagtg agttaactgag ctttggtcca ctaaaaaag 2340
ctgagctttg gtccactaaa acaagatgaa aaatacaaga gtgactctat aactctgggtc 2400
tttaagaaag ctgccttttc attttttagac aaaatctttt caacgctgaa atgtacctaa 2460
tctggttcta ctaccataat gtatatgcag cttcccaggag atgaatgctg tgtttaaatt 2520
tcataaagta aatttgtcac tctagcattt tgaatgaata gtcttcaact tttaaattat 2580
tcatcttctc tataataatg acatcccagt tcatggaggc aaaaaacaag tttcttgtaa 2640
tcctgaaact ttctatgctc agtggaaagt atctgccagc cacagcatga ggcctgtgaa 2700
ggctgactga gaaatcctct gctgaagacc cctgggtctg ttctgcctcc aacatgtata 2760
attttatttg aaatacataa tcttttcaact atgcttttgt ggggtttttt ttaagtatgt 2820
gtaaaaatgt gatgctcaga taagtacatt tatatcagtt cagtgttaaa atgcagtctc 2880
ttgagttaaa gtcattctta ttttaaattgc agtgataaat gtcaactcct cggagaaact 2940
aggagaacaa caacagaaag ctgtgtttgt cttttttctc tcaaatatat ctcccgtatg 3000
agatttcagg tccccatgtt ttcaccaagc aatctgctat gtcagccaac ccaacatcac 3060
tttctacagg aggttatgat ttttgccatt tactagagga agatgtttta tgaaatcaat 3120
ttgggggttg aattcagggt cagtcatcag ttctttaggg gctgcaatgt tttaaaaaaa 3180
ataagtcac agatttttaag aaaaaagtga tgatttctta ttgatatttt tgtaacagaa 3240
tatagctctt aactgaaaat ccagaaccag aaacataaat cttgagtttc ttttcatgta 3300
cataaaaaagc aatagccttt tagtatagat agccctgagc caaaaagtaa tagaattttc 3360
tctagatatt taatacacag agtgatataga ctgactctaa gttaataatg tgcaaaatat 3420
cttaaacatc cctcccctta ttcaacaatt atgtatcagt gatcttgaac cattgtttta 3480
tatttttcac ctttgtaacc tcatggaaag aggcctttaca tactttctat gtactattta 3540
cttagaaggg agcccccttc cagtcatgaa acttcatttg ttttatocat atccctgagg 3600
actgtgtaga ctttatgtca gttctgtgta gactttatgt cagtttttgt cattatttga 3660
aaatctattc tgacaacttt ttaattcctt tgatcttata agttaagct gtaacaactg 3720
aaattgcatg gatcaagtaa gcatagtttt atccaggagg aaaaataaaa ggaagccata 3780
gaattgctct ggtcaaaaacc aagcacacca tagccttaac tgaatattta ggaaatctgc 3840
ctaattctgt tatatttggt gtttggtttt tgactgttgg gctttgggaa gatgttattt 3900
atgaccaata tctgccagta acgctgttta tctcacttgc tttgaaagcc aatgggggaa 3960
aaaaatccat gaaaaaaaaa agattgataa agtagatgat tttgtttgta tccctacca 4020
tctcctggca gccctactga gtgaaattgg gatacatttg gctgtcagaa attataccga 4080
gtctactggg tataacatgt ctcaattgga aagctagtag ttttaaatgg gtgccaaagg 4140
tcaactgtaa tgagataatt atccctgcct gtgtccatgt cagactttga gctgatcctg 4200
aataataaag ccttttacct t 4221

```

<210> 402

<211> 3224

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018401

<400> 402

```

gtcccacatc ccgcatccgg catcccagcg gccgggcatg tagcagcggc agcaacggcg 60
gaatatgggc gggaaccact cccacaagcc ccccggtgtt gacgagaatg aggaagtcaa 120
ctttgaccat ttccagattc tgcgggccat tggtaaaggg agttttggaa aggtatgcat 180
cgtgcagaag cgagacacta agaaaatgta tgcaatgaag tacatgaaca agcagaagtg 240
catcgagagg gatgaggttc ggaatgtttt ccgggagctg cagatcatgc aagggtctgga 300
gcaccccttc ctgggtcaatc tgtgtgtact cttccaggat gaggaggaca tgttcatggt 360
ggtggacctg ctccctgggag gcgacctgcg ctaccatctg cagcagaatg tgcatttcac 420

```



```

agaggggact gtgaaactct acatctgtga gctggcactg gccctggagt atcttcagag 480
gtaccacatc atccacagag acatcaagcc agacaatatc ctgctggatg aacacggaca 540
tgttcacatt acagacttca acatagcgac ggtagtgaaa ggagcagaaa gggcttcctc 600
catggctggc accaagccct acatggctcc agaagtattc cagggtgtaca tggacagagg 660
ccccggatac tcgtaccctg tcgactgggtg gtccctgggc atcacagcct atgagctgct 720
gcggggctgg aggcgttacg aaatccactc ggtcacgccc atcgatgaaa tcctcaacat 780
gttcaagggtg gagcgtgtcc actactcctc cacgtgggtg aaggggatgg tggccctgct 840
gaggaagctc ctgaccaagg atcctgagag ccgctgtgct agccttcatg acatacagag 900
cgtgccctac ttggccgaca tgaactggga cgcggtgttc aagaaggcac tgatgcccgg 960
ctttgtgccc aataaaggga ggttgaactg cgatcccaca tttagcttg aagagatgat 1020
tctagaatcc aagccacttc acaaaaagaa gaagcgattg gcaaagaaca gatccaggga 1080
tggcacaag gacagctgcc cgctgaatgg acacctgcag cactgtttgg agactgtccg 1140
ggaggaattc atcatattca acagagagaa gctcaggagg cagcaggggac agggcagcca 1200
gctcttgga accgacagcc gagggggagg ccaggcccaa agcaagctcc aggacgggtg 1260
caacaacaac ctccctaccc acacctgcac ccgtggctgc agcagctgag cccacacttg 1320
ttgctgctca acaggactgc actcgtctct gccctgcccc cccagagccc ctctttgtgc 1380
cctgatggtc cctgtctcac ccctgaaaac atcagatgca gaaaaagccc tggacttgga 1440
gtgggaagc ctgggttctg gtcccatctc catgactgat tcacgtgtga cctcagacaa 1500
gtcacgccct ctctgtgcct ccgttttctg catctgccaa aggggttaaa cacttctgcc 1560
ccacttcaaa ttacaagatt atggggagaa cccaattagg taggaaacat gaaaaacctt 1620
tgatatttat aaaatcattt ttacgtgcaa aatataacct taatatttga agtgaccccc 1680
attcccaaaa gcaatcaaac cgtcatgact ttgcaatttg gcacatccta gcttgtaga 1740
gggcacttcc gaaaaacaca gccctgacag caaaataaag gtctgatatg ttggcccctt 1800
ctatggaaac aacgctgcc aatcctggag caaaacctga agtgtcttca tgtgcattct 1860
ctggcaggcc acagtcctga gcttgaaga ttgtgcagca tgcagaccag acttgtcccc 1920
aagggtcag cgctgcggtc tcaactcctc cctcatttaa gaagactatc cttacctttt 1980
agtttcagca gtccctacca ccaccatctc ccagtgctg ggatggcaca cagggtgtcca 2040
ttcagatgag agttgggtcg ctgagcattg gttactcctg cagagtgtaa tcagaccccc 2100
atccaactgg ccgaaagcc cagacctgca gcagaactct ccaactctct atcagctttc 2160
agggttttct ctccctggaa ggggtgtaaaa tcagcttgtc agattcttct tacagagagt 2220
atccaatcgg tattgggtgga gcggctccct atttatacaa taggaagcat ggggtgcttag 2280
aaagtttatt tcaggaggaa aatgggttca cacaaaaagc aaactacatt ctgatctgct 2340
cagggagaa ctggcctttg aactggaaga ttgtgggatg agcagggaaa gcttagactt 2400
tggagtccag tttgtgttca gaatccagcc ctgctggcta ctaactaact gggagacctt 2460
aggcaaagca tgcaatcgct ctgaatggca gtttctcat ttttaaacag ggataataaa 2520
actaatattg caggggagtt acaggggttaa ataagatcct gtgtgtaacc ccaagcattg 2580
gatgactcat agaatggcct tttttgtcag cataatcgct atcattattt agatactttc 2640
ttccttctact caccagcag gtcagttttc tgtgcaaaca aacctgttta ggattcttcc 2700
aaatgttctt cctgggtctt ttgatatttg ttgtttacat cctgctgaag ttcgactgtg 2760
tttttatttt ttcatccaac ttccattttt cactttttac atgattactc aatccttggg 2820
gctgtccatg tcatctctta gatctcttaa aagacatttt aatgtatggt taggttttat 2880
atttttattt tttaaaaaag aaatagtcag tgttttctc ctttcaaccg agactatttc 2940
tggatttgtg gtcctcgctc agttgacttg ttttgacac ttttctttac ttcatgtccc 3000
catcaacaac cgtcctgctc cccacctccc ccaggaaata aggggcctgc tcctctccct 3060
actgtgaccc tggaggctct taagatgatg atggtttttt ttattgggct gagttcacga 3120
attaggggca ggagctggaa gtcgccctag gaacaccaga tttcctggtt ctgttcaagt 3180
tggcatttct tgtttggaat aaactatttc ttggacatto cttc 3224

```

<210> 403

<211> 2460

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000788

<400> 403

```

aaagtcaaac cccgacaccg cggcggggcg gtgagctcac tagctgaccc ggcagggtcag 60
gatctggcct agcggcgccg cgagctccag tgcgcgcacc cgtggccgcc tcccagccct 120
ctttgccgga cgagctctgg gccgccacaa gactaaggaa tggccacccc gcccaagaga 180
agctgcocgt ctttctcagc cagctctgag gggacccgca tcaagaaaat ctccatcgaa 240

```

```

gggaacatcg ctgcaggga gtcaacattt gtgaatatcc ttaaacaatt gtgtgaagat 300
tgggaagtgg ttcttgaacc tgttgccaga tgggtgcaatg ttcaaagtac tcaagatgaa 360
tttgaggaac ttacaatgtc tcagaaaaat ggtgggaatg ttcttcagat gatgtatgag 420
aaacctgaac gatgggtctt taccttccaa acatatgcct gtctcagtcg aataagagct 480
cagcttgccct ctctgaatgg caagctcaaa gatgcagaga aacctgtatt attttttgaa 540
cgatctgtgt atagtgcag gtatatTTTT gcatctaatt tgtatgaatc tgaatgcag 600
aatgagacag agtggacaat ttatcaagac tggcatgact ggatgaataa ccaatttggc 660
caaaccttg aattggatgg aatcatttat ctccaagcca ctccagagac atgcttacat 720
agaatatatt tacggggaag aaatgaagag caaggcattc ctcttgaata tttagagaag 780
cttcattata aacatgaaag ctggctctcg cataggacac tgaaaaccaa cttcgattat 840
cttcaagagg tgcctgtctt aacactggat gttaatgaag actttaaaga caaatatgaa 900
agtctgttg aaaaggtcaa agagttttgg agtactttgt gatcttgctg aagactacag 960
gcagccaaat ggttcagat acttcagctt tgtgtatctt cgtaacttca tattaatata 1020
agtttcttta gaaaaccaa gtttttaatc gtttttgttt taaggaaaaa agatttttaa 1080
aatgaatctt atgcaaaact ttttgatcag tttcttttct tttgtttttt ttttaaaaaa 1140
gacattttaa gacaaagaca ttatttctca tagcaggaaa tgtagaggta gatggttcca 1200
gtatcagcat agtgactaaa ctacattata aaagatccag ctctcttctg tcattccctc 1260
cttttgtctt cctcagcagg ttggcttttt tccttgggtg ctctcacttc gttggtgacc 1320
agtttcttaa actgaaagct ttaattgtac atagtaaatg gtatgtgtgc ctgtgtaaat 1380
tagtgtacct attaaaagtt gcaaagtggg attaaaggaa tccctagaat aaggattctg 1440
aagttttatt ttaaatatt atcttcttaa cagtttagtc ccacctctta ctctctgct 1500
cagctgtctt tctctactgt ctggattaat taggcagcct gctataaagt taaagtcaca 1560
catttctatt ttgcaaacac tgtgattact ctgtgctttg tagtttgctt tgctttgtag 1620
ggttctgctt ttaagttttt ctctttttca gacaaattac tgataaaaaa gatattgctc 1680
tatatgtaat atatcctgaa agcattattt tttgttgaat aggaaataaa attaatgaag 1740
acagaggcta gaaagcatcc attaatatt gagacacact taactactta tctctaaacc 1800
atctatgtga atatttgtaa aaataatgaa tggactcctc ttagttctgt atataaatat 1860
attttctttc tagtttgttt agttaagggtg tgcagtgttt ttctgtgtg ttaaaccctt 1920
ccattttacg ttttagaaaa ttttatgtat tttaaaataa ggggaagagt cattttcacc 1980
tttaaactac tatttttctt tccaagtcac ttttggtttt ggttctttat tcaaagatga 2040
taatttagtg gattaaccag tccagacgca ctgatctttg caaaggagac ttaatttcaa 2100
atctgtaatt accatacata aactgtctca ttatacgtat gcattttttt agtttgtttt 2160
tgtttgggtat aaataattt gttaatataa ttttcttaa gtataaacct tatgaactac 2220
agtggagcta cactcattga aatgtaattt cagttctaaa aagatgtaat aatcatttta 2280
gaattaaaat ttattctact tttaaataaa ttatgaatat taaaggtgaa aattgtataa 2340
attactttga ttccatttta agtggagaca tatttcagtg atttttagta accttataaa 2400
atgtataatg acttttaaaa tttgtagaat tgaaaagacg ctaataaaaaa tttattattt 2460

```

<210> 404
<211> 558
<212> DNA
<213> Homo sapiens

<300>
<308> Contig51464_RC

```

<400> 404
agagttttcac aaataacatt tattgaccac attcataaca gaaactcacc cacattcaca 60
tgaacttttgc tacaaccaag aatcacctga tgagaaatca tttgggttga atcctgggtg 120
tggtattctgt tctgagaaga cagcgactgt cacagacgcc tcctgattgg tggactctca 180
aagatcttga tcgtgacaca tttttgtggc tgttcttgac agtagctgaa gattccaact 240
tcaaacccca aaacactggc gaatggccat gatgaccctc gggatggcct ggaggagggt 300
caggaggccg aggggcagca gacgcgctgc aggagccggt agctggctct ggtagggggc 360
ctggagctgc tggaaacttg acaacaggca ctcatcaccg cagatccgcc agcagttgtg 420
ccagcagcat gcattccagca ggggtgctgc tttcggggtc tcagggacag agggaccag 480
taaggcagtg ctgcttccac cctggacaca ggcaggcggc ccccgatctt ggtctggcat 540
tcacagcaga tggccagg
558

```

<210> 405
<211> 1929

<212> DNA
 <213> Homo sapiens

<300>
 <308> AL080079

<400> 405
 gggggagaat ttccattagc caagctgtat gcagtggaaat tttcctttta ggagacacac 60
 aattaagact ctctggttct gtccttgctt tggagacttt aagacatttc cttaaagcaca 120
 aataaaagcc tcgtatttcc ccattgagag ttttgttcca aggaatatga agtgagacat 180
 atgggtgagt tataataatc aaaataattt atgaagagct gggctctgcaa tagctagtct 240
 aaaaactact tgtgtgtcag tcctctggtt atagtatata agagcctgag gaggtctggc 300
 aagatagatg gtgtattatt tatggatcag gctgctgcat acaaaccttg catactatta 360
 tgcagcttac ctaactctca gactattctg agtaatgctt gcttgctaataaatcggtc 420
 gagaccacat tgtaattggt cttagatgat ggagtcocat cagtttctta gaaatcggtc 480
 tcagtgcagt ctgtgctttt tcacatttgc tctgggttat ctgggaagta tcaggttctg 540
 ggaggcaaca gcattaagtgc ataagaaaag gagacattct ggcaaagcca atctgcttaa 600
 aggcaaagtc cagaacctgg aacctagagg cctttctctc tgcacgaaaa acaggtagtt 660
 tgcagtctga gatattggag agcttttagg ctacacagca acccaagggg cctctcacct 720
 tttgctgagc ttcaatcagg aagctatttg cctggttcca gcagatgat agataatgag 780
 gtagtggggt ttttattact gttccatttt gcaacatcct gcaacacat cctgggagac 840
 aagagcatta ccgagcttgg ctttcacggg ggagggttgt attcagtaaa aaagaatagt 900
 aaatataagg tcaactgagat tctaagtaag atagtaaata taaggtcact gagccaaatc 960
 cttttcaata ggcatatatt aacaggctgc ttattttggc ggaggttaca tatggatgaa 1020
 aatgaatctt agtcactgaa tattcatata ctttcccc aaacacctag acattcatag 1080
 tagattttta ttagttagct ttctaactag tcagatttct gcccaaagtg cttagtcaat 1140
 agtaattaag atatagcttc aaaaatggct ttgcttttga tttctactca tttcgtatg 1200
 gctccagaaa aataattttt tcatatttga caatgtcagc tccactttag aaattttcaa 1260
 taaccagatg agaaaaaaat taagaaattg ctcaagggaa acatttgtaa atggatttga 1320
 aagattgagc caaattctgt tgtcagttct aagcatgcag ttctcacctc catttagtcc 1380
 cccatcagaa cagaggctcag gaatttagct ggggagccta aatttagttc agcttacctt 1440
 tgagaatagc atcaattcag actctctttt cattatgttt tcttttcttt ttccctcttt 1500
 ttaactaca ttgtgttaga gtcgtagtct aggatcctga gagattttcc attctgttca 1560
 ccattcacct gcattgtaaa gattttcttt gtctgttgtt ggcatagatt cttttgtaca 1620
 tattttattt tttgtgttta tatatgtcaa ttggtttcct ttcttagctt gatattgcct 1680
 agctttgttg ttttaattaa ctttctatta gagggactgt atatattttt tctaaatact 1740
 ttgtgaaatc atttttggtg gcaatatctt tgaatatgat gaataaaagt gactgtgagt 1800
 gcaaatagaa tttagcagtaa gaagctactc tagctaattt gccatttcac ttaaatggaa 1860
 aatgtttttc aaataaatac ttatgttgtt catgttccaa gttattcaaa aaaaaaaaaa 1920
 aaaaaaac 1929

<210> 406
 <211> 3915
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_006931

<400> 406
 gtgggggtgg gtggggctgg gggcttgcgc ccctttcagg ctccaccctt tgcggagatt 60
 ataaatagtc atgatccag cgagaccag agatgcctgt aatggtgaga ctttggatcc 120
 ttctgagga cgtggagaaa actttctgct gagaaggaca ttttgaagg tttgttggct 180
 gaaaaagctg tttctggaat caccctaga tctttcttga agacttgaat tagattacag 240
 cgatggggac acagaaggtc accccagctc tgatatttgc catcacagtt gctacaatcg 300
 gctctttcca atttggctac aacactgggg tcatcaatgc tcctgagaag atcataaagg 360
 aatttatcaa taaaactttg acggacaagg gaaatgcccc accctctgag gtgctgctca 420
 cgtctctctg gtccttgcct gtggccatat tttccgtcgg gggatgatgc ggctcctttt 480
 ccgtcggact cttcgtcaac cgctttggca ggcgcaatc aatgctgatt gtcaacctgt 540
 tggctgtcac tgggtgctgc tttatgggac tgtgtaaagt agctaagtcg gttgaaatgc 600
 tgatcctggg tcgcttgggt attggcctct tctgcggact ctgcacagg tttgtgcca 660

```

tgtacattgg agagatctcg cctactgccc tgcgggggtgc ctttggcact ctcaaccagc 720
tgggcatcgt tgttgggaatt ctgggtggccc agatctttgg tctggaattc atccttgggt 780
ctgaagagct atggcgcgtg ctactgggtt ttaccatcct tctgctatc ctacaaagt 840
cagcccttcc attttgccct gaaagtccca gatttttgct cattaacaga aaagaagagg 900
agaatgctaa gcagatcctc cagcgggtgt ggggcaccca ggatgtatcc caagacatcc 960
aggagatgaa agatgagagt gcaaggatgt cacaagaaaa gcaagtcacc gtgctagagc 1020
tcttttagagt gtccagctac cgacagccca tcatcatttc cattgtgctc cagctctctc 1080
agcagctctc tgggatcaat gctgtgttct attactcaac aggaatcttc aaggatgcag 1140
gtgttcaaga gcccatctat gccaccatcg gcgogggtgt ggtaataact atcttctactg 1200
tagtttctct atttctggtg gaaagggcag gaagaaggac tctgcatatg ataggccttg 1260
gagggatggc ttttgttcc acgctcatga ctgtttcttt gttattaaag gataactata 1320
atgggatgag ctttgtctgt attggggcta tcttggtctt tgtagccttc tttgaaattg 1380
gaccaggccc cattccctgg tttattgttg ccgaactctt cagccagggc ccccgcccag 1440
ctgcgatggc agtggccggc tgctccaact ggacctccaa ctctctagtc ggattgctct 1500
tcccctccgc tgctcactat ttaggagcct acgtttttat tatcttcacc ggcttctctca 1560
ttaccttctt ggcttttacc ttcttcaaag tccctgagac ccgtggcagg acttttgagg 1620
atatcacacg ggcccttgaa gggcaggcac acggtgcaga tagatctgga aaggacggcg 1680
tcatggagat gaacagcatc gagcctgcta aggagaccac caccaatgtc taagtctgtc 1740
ctccttccac ctccctcccg gcatgggaaa gccacctctc cctcaacaag ggagagacct 1800
catcaggatg aaccaggac gcttctgaat gctgctactt aattcctttc tcatcccacg 1860
cactccatga gcaccccaag gctgcggtt gttggatctt caatggcttt ttaaatttta 1920
tttctggac atctcttct gcttaggaga gaccgagtga acctaccttc atttcaggag 1980
ggattggccg cttggcacat gacaactttg ccagcttttc ctcccttggg ttctgatatt 2040
gccgcaactag gggatatagg agagggaaaag taagggtgcag ttccccaac ctgagactta 2100
ccaggaagca gatacatatg agtgtggaag ccggagggtg tttatgtaag agcaccttc 2160
tcacttccat acagctctac gtggcaaatt aacttgagtt ttatttattt tatcctctgg 2220
tttaattaca taattttttt ttttttactt taagtttcag gatacatgtg ccgaatgtgc 2280
aggtttgta cataggtata tatatgccat gatggaaata tttatttttt taagcgtaat 2340
tttgccaaat aataaaaaca gaaggaaatt gagattagag ggaggtgttt aaagagaggt 2400
tatagagtag aagatttgat gctggagagg ttaagggtca ataagaattt agggagaaat 2460
gttgttcatt attggagggt aaatgatgtg gtgcctgagg tctgtacgtt acctcttaac 2520
aatttctgtc cttcagatgg aaactcttta acttctcgta aaagtcatat acctatataa 2580
taaagctact gatttccctg gagctttttt ctttaagata atagtttaca tgtagtagta 2640
cttgaaatct aggtatttta actaatatgg gcattgtagt taatgatggt tgatgggttc 2700
taattttgga tggagtccag ggaagagaaa gtgatttcta gaaagcctgt tcccctcact 2760
ggatgaaata actccttctt gtagtagtct cattactttt gaagtaatcc cgccacctat 2820
ctcgtgggag agccatccaa ataagaaacc taaaataatt ggttcttggg agagattcat 2880
tatttttcca ctttgttctt taggagattt taggtgttga ttttctgttg tattttaact 2940
cataccttta aaggaattcc ccaaagaatg tttatagcaa acttggaaat tgtaacctca 3000
gctctgggag aggtattttt tctgagcgat tattatctaa agtgtgtgtg tgctttaggc 3060
tcacggcacg cttgcgtatg tctgttacca tgtcactgtg gtcctatgcc gaatgccctc 3120
aggggacttg aatctttcca ataaaccagg tttagacagt atgagtcaat gtgcagtgtg 3180
gccacactt gagaggatga atgtatgtgc actgtcactt tgctctgggt ggaagtacgt 3240
tattgttgac ttattttctc tgtgtttgtt cctacagccc ctttttcata tgttgctcag 3300
tctcccttcc ctttcttggg gcttacacat ctgagaccct ttagccaaac ccttgctcag 3360
gacagtattt tgggtcttag ttctcactgt tccctctgct cctggagcct ttgaataaaa 3420
atgcacgtag ctgaggccgg atgcgggtgc tcacgcctgt aatcccagca ctttgggagg 3480
cctaggcggg cggtcagggg ttccgagacca gtctggccaa catcgtgaaa ccctgtctct 3540
actaaaaatg caaaaattag ccgggcgttg tggcgggcgc ctgtaatccc agctacttg 3600
gaagctgagg cgggagaatc atgtgaaccc gggacgcagg ggttgcaagt agcggagatc 3660
gcatcattgc actctagcct gggccacagg gcgagactcc gtctcaaaaa aaaaaaatg 3720
cacatagcta tcgagtgtgc tttagcttga aaaggtgacc ttgcaacttc atgtcaactt 3780
tctggctcct caaacagtag gttggcagta aggcagggtc ccatttctca ctgagaagat 3840
tgtgaatatt tccatatgga ttttctattg ttactctggg tctttgtttt aaaataaaaa 3900
ttctgaatgt. acacg 3915

```

<210> 407

<211> 2125

<212> DNA

<213> Homo sapiens

<300>

<308> AF257157

<400> 407

```

tggcgatggc gtacttggct tggagactgg cgcgccgcttc gtgtccgagg tcactagttt 60
cccggtagtt cagctgcaca tgaatagaac agcaatgaga gccagtcaga aggactttga 120
aaattcaatg aatcaagtga aactcttgaa aaaggatcca ggaaacgaag tgaagctaaa 180
actctacgcg ctatataagc aggccactga aggacottgt aacatgcccc aaccagggtg 240
atltgacttg atcaacaagg ccaaatggga cgcattggaat gcccttggca gcctgccccaa 300
ggaagctgcc aggcagaact atgtggattt ggtgtccagt ttgagtcctt cattggaatc 360
ctctagtcag gtggagcctg gaacagacag gaaatcaact ggttttgaaa ctctggtggg 420
gacctccgaa gatggcatca caaagatcat gttcaaccgg cccaaaaaga aaaatgccat 480
acacactgag atgtatcatg aaattatgct tgcacttaaa gctgccagca aggatgactc 540
aatcatcact gttttaacag gaaatggtga ctattacagt agtgggaatg atctgactaa 600
cttcaactgat attccccctg gtggagtaga ggagaaagct aaaaaaatg ccgttttact 660
gaggtaagag ttactgacct ttgtctgac ccagagggcc cgttcttagt catcaaagtg 720
gtatttagtt gagcaaaagt gcaatatctg cttctctgtg tcagtgcagt taattgtgac 780
tttgacaag catatttctc taaactggct caatccatag gtgggcattc attctttctc 840
tttgacctct gttacctgtt catcctgctg gtctcatgag atttgactgg aaacataaaa 900
agtgaactat agttgtaaac aaaattatct tgcattgagg cagaggaatt tagcaaatc 960
tcatccagcc agaggaatgg gagatttcca tgaacgctga aggtaggaga ctcgaggagg 1020

```

```

tcttgtaaac acaggaggag gtcatgaagg cttttctott tttcctgcc aaatctgtaa 1080
ttcccacttg ctgtgatatt gcagcatgtg ctacgtcacc caacctaatc tctaagactg 1140
cggtctggtt acagccctga actcactgtc cagaatacaa gattcaaacc aggggttgacc 1200
cataggtctc actggttggg cttccatcca ctgccattca gacggatgcc ctggatgaat 1260
ctggtgaggga agaccaggga agtggcagga caaggaggca ctgctggtgg tggctgtgga 1320
aacaccacag ctgacagaag gctttatctc agagtacagc tgtcttgccc atgcttcagt 1380
catgctgctt tggcacattt aatgattttg tgtttccatt tctgtaggga atttgtgggc 1440
tggtttatag attttcttaa gcctctgatt gcagtggtea atggtccagc tgtgggcac 1500
tccgtcacc ccttgggct attcgatgcc gtgtatgcat ctgacagggc aacatttcat 1560
acaccattta gtcacctagg ccaaagtcog gaaggatgct cctcttacac ttttccgaag 1620
ataatgagcc cagccaaggc aacagagatg cttatttttg gaaagaagtt aacagcggga 1680
gaggcatgtg ctcaaggact tgttactgaa gttttccctg atagcacttt tcagaaagaa 1740
gtctggacca ggctgaaggc atttgcaaa cttcccccaa atgtcttgag aatttcaaaa 1800
gaggtaatca ggaaaagaga gagagaaaaa ctacacgctg ttaatgctga agaatgcaat 1860
gtccttcagg gaagatggct atcagatgaa tgcacaaatg ctgtggtgaa cttcttatcc 1920
agaaaatcaa aactgtgatg accactacag cagagtaaag catgtccaag gaaggatgtg 1980
ctgttacctc tgatttccag tactggaact aaataagctt cattgtgcct tttgtagtgc 2040
tagaatatca attacaatga tgatatttca ctacagctct gatgaataaa aagttttgtg 2100
aaccaaaaaa aaaaaaaaaa aaaaaa 2125

```

<210> 408

<211> 1446

<212> DNA

<213> Homo sapiens

<300>

<308> NM_002916

<400> 408

```

ctgccattta ggacaagctg gatgatgatg gtttgatagc tccagggggt cgtgtatagg 60
agatgatgaa tctgcttcat ccagaatcac aatcttaaaa ggcgggaact gaggcgactg 120
tggggacatc agtgatcgta agtctcctgg gcccggtatt ctacagattag gtgacggagc 180
taagacttct agaccatctc gtcctttttg tatcgcgga acctgaggaa cgagccggcg 240
gcggtgacct gcacgagaag ccaggctaac tgggtgaagt accatgcaag catttcttaa 300
aggtacatcc atcagtacta aacccccgct gaccaaggat cgaggagtag ctgccagtgc 360
gggaagtagc ggagagaaca agaaagccaa acccgttccc tgggtggaaa aatatcgccc 420
aaaatgtgtg gatgaagtgt ctttccagga agaagtgggt gcagtgtgta aaaaatcttt 480
agaaggagca gatcttcta atctctgtt ttacggacca cctggaactg gaaaaacatc 540
cactattttg gcagcagcta gagaactctt tgggcctgaa ctttccgat taagagttct 600

```

```

tgagttaaatt gcatctgatg aacgtggaat acaagtagtt cgagagaaag tgaaaaattt 660
tgctcaatta actgtgtcag gaagtcgctc agatgggaag ccgtgtccgc cttttaagat 720
tgtgattctg gatgaagcag attctatgac ctacagctgct caggcagctt taagacgtac 780
catggagaag gagtcgaaaa ccaccgatt ctgtcttata tgtaactatg tcagtcgaat 840
aattgaaccc ctgacctcta gatgttcaaa attccgcttc aagcctctgt cagataaaat 900
tcaacagcag cgattactag acattgccaa gaaggaaaat gtcaaaatta gtgatgaggg 960
aatagcttat cttgttaaag tgtcagaagg agacttaaga aaagccatta catttccttca 1020
aagcgctact cgattaacag gtggaaagga gatcacagag aaagtgatta cagacattgc 1080
tggggtaata ccagctgaga aaattgatgg agtatttgct gcctgtcaga gtggctcttt 1140
tgacaaacta gaagctgtgg tcaaggattt aatagatgag ggatcatgcag caactcagct 1200
cgtcaatcaa ctccatgatg tggttgtaga aaataactta tctgataaac agaagtctat 1260
tatcacagaa aaacttgccg aagttgacaa atgcctagca gatgggtgctg atgaacattt 1320
gcaactcatc agcctttgtg caactgtgat gcagcagtta tctcagaatt gttaacgtga 1380
tatatctgga tgggggggtt tgtaaataat gaagttgtaa taaaaataaa atgacccaaa 1440
gcaccg 1446

```

<210> 409

<211> 788

<212> DNA

<213> Homo sapiens

<300>

<308> Contig55725_RC

<400> 409

```

caagttagaa caaacatttt attattaaaa taaacttttg tataaaagca ttacagatca 60
aaagctgtat ttacacttat cgattcaagg tccaattatg catcaaacat tgaatggcac 120
agcaatgggt tacatatgca agtaaattgg acatacaaac acttagattc cacctctacc 180
aaataccttg attaattgca agaggagggg gaatactgac acaggaaaac tgcccagaaa 240
ctagactggc agagatgtca ggtaaacaaa ctgctaaaag ttacatctcc aaaaaggcac 300
ttatcattgt tataaaagtg cttaaaatct aaacttgaac cttgtgcttg gtttataaat 360
ttacaagaaa ctgcaaagaa ccacagacta gtttttaata tcaagtttcc atacaaaatt 420
gtccaagaat ttatttgcaa taccttacct gtgaactgaa ataaaacttg caaactacaa 480
ttatactaaa gttttatact gaaaattcaa atcagcttac tcataatctc cagatagctc 540
ttccatcatt attttaatta agtattacat tgcccaccaa cctttgaatt atgatcgtgt 600
aagctggacg caccacttac ccttcgggca gaagaaatta ggagaaatta agtctcaact 660
gtgacaacaa cttgtaacat aagcgcaata catgggattg agagacatca atctgcaata 720
attacctaatt ttttccacat ggttttaaga cttttgtgac ttgattctat taaaaatttc 780
aactagta 788

```

<210> 410

<211> 433

<212> DNA

<213> Homo sapiens

<300>

<308> Contig24252_RC

<400> 410

```

cacgtaacac aaaagctttt acttggaac tggcaaatac tggactagaa tactgacatg 60
ctcacgctct ggcgggagct cggatgcaga agctattgca caaagcccct ctgattgcct 120
tgctcctctt tggcatgtat cagcagagcc ccaagggccca attgccaca ggtggggact 180
gttctccatc aaggtatggg gaccctact tccttgttt gttaaaaagt gcaggtaggc 240
gaagaagccc aggcagttga ccagctctt gaaacagctg actcccagc acccagtgtc 300
cactggcaga tgagaggttt gcacctatac cagccctgtt tccttggtta gggctcaagt 360
tgtcagtgtc caggtacctg ggctgggcca gagctggggc caggggagcc tggctcttgag 420
gggtaagggc tgc 433

```

<210> 411

<211> 969

<212> DNA

<213> Homo sapiens

<300>

<308> AF201951

<400> 411

```

acggaacagc cagtgggagg ttccagctga gcgctcccca gaggtgagct gatccccagc 60
cacagcacac aggaccaggc tgcgagaaca gcatcatcag catcatgcta ttacaatccc 120
aaacctgggg gggtttctcac agcttttacac caaaggggcat cactatccct caaagagaga 180
aacctggaca catgtaccaa aacgaagatt acctgcagaa cgggctgcca acagaaacca 240
ccgttcttgg gactgtccag atcctgtgtt gcctgttgat ttcaagtctg ggggccatct 300
tggtttttgc tccctacccc tcccacttca atccagcaat ttccaccact ttgatgtctg 360
ggtacccatt tttaggagct ctgtgttttg gcattactgg atccctctca attatctctg 420
gaaaacaatc aactaagccc tttgacctga gcagcttgac ctcaaagca gtgagttctg 480
ttactgcagg agcaggcctc ttccctcctg ctgacagcat ggtagccctg aggactgcct 540
ctcaacattg tggctcagaa atggattatc tatcctcatt gccttattcg gactactatt 600
atccaatata tgaaatcaaa gattgtctcc tgaccagtgt cagtttaaca ggtgtcctag 660
tggtgatgct catcttcact gtgctggagc tcttattagc tgcatacagt tctgtctttt 720
ggtggaaaca gctctactcc aacaaccctg ggagttcatt ttccctcgacc cagtcacaag 780
atcatatcca acaggtcaaa aagagttctt cacggtcttg gatataagta actcttggcc 840
tcagaggaag gaaaagcaac tcaacactca tggccaagtg tgattagact ttccctgaaat 900
ctctgccatt ttagatactg tgaacaaaac taaaaaaaaa aaagcttttg ttttgtaaaa 960
aaaaaaaaa

```

<210> 412

<211> 2917

<212> DNA

<213> Homo sapiens

<300>

<308> NM_005915

<400> 412

```

ggaaaaaagc gacttgtggc ggtcgagcgt ggcgagggcg aatcctcggc actaagcaaa 60
tatggacctc gcggcgggcag cggagcgggg cgccggcagc cagcacctgg aggtccgcga 120
cgaggtggcc gagaagtgcc agaaactggt cctggacttc ttggaggagt ttccagagcag 180
cgatggagaa attaaatact tgcaattagc agaggaaactg attcgtcctg agagaaacac 240
attggttgtg agtttttgtg acctggaaca atttaaccag caactttcca ccaccattca 300
agaggagtgc tatagagttt acccttacct gtgtcgggcc ttgaaaacat tcgtcaaaga 360
ccgtaaagag atccctcttg ccaaggattt ttatgttgca ttccaagacc tgccctaccag 420
acacaagatt cgagagctca cctcatccag aattggtttg ctcaactcgca tcagtgggca 480
ggtggtgctg actcaaccag ttcaaccaga gcttgtgagc ggaacttttc tgtgcttgga 540
ctgtcagaca gtgatcaggg atgtagaaca gcagttcaaa tacacacagc caaacatctg 600
ccgaaatcca gtttgtgcca acaggaggag attccttactg gatacaaata aatcaagatt 660
tgttgatttt caaaagggtc gtattcaaga gacccaagct gagcttcctc gagggagtat 720
ccccgcagc ttagaagtaa ttttaagggc tgaagctgtg gaatcagctc aagctgggtga 780
caagtgtgac tttacaggga cactgattgt tgtgcctgac gtctccaagc ttagcacacc 840
aggagcacgt gcagaaacta attcccgtgt cagtgggtgt gatggatatg agacagaagg 900
cattcgagga ctccggggccc ttggtgttag ggacctttct tataggctgg tctttcttgc 960
ctgctgtgtt gcgccaaacca acccaagggt tggggggaaa gagctcagag atgaggaaca 1020
gacagctgag agcatthaaga accaaatgac tgtgaaagaa tgggagaaag tgtttgagat 1080
gagtcaagat aaaaatctat accacaatct ttgtaccagc ctgttcccta ctatacatgg 1140
caatgatgaa gtaaaacggg gtgtcctgct gatgctcttt ggtggcggtc caaagacaac 1200
aggagaaggg acctctcttc gaggggacat aaatgtttgc attgttgggtg acccaagtac 1260
agctaagagc caattttctca agcacgtgga ggagttcagc cccagagctg tctacaccag 1320
tggtaaagcg tccagtgtct ctggcttaac agcagctgtt gtgagagatg aagaatctca 1380
tgagtttgtc attgaggctg gagctttgat gttggctgat aatgggtgtg gttgtattga 1440
tgaatttgat aagatggacg tgcgggatca agttgctatt catgaagcta tggaaacagca 1500
gaccatatcc atcactaaag caggagtga ggctactctg aacgcccga cgtccatttt 1560
ggcagcagca aacccaatca gtggacacta tgacagatca aaatcattga aacagaatat 1620
aaatttgtca gctcccatca tgtcccgatt cgatctcttc tttatccttg tggatgaatg 1680

```

taatgaggtt	acagattatg	ccattgccag	gcgcatagta	gatttgcatt	caagaattga	1740
ggaatcaatt	gatcgtgtct	attccctcga	tgatatcaga	agatatcttc	tctttgcaag	1800
acagtttaaa	cccaagattt	ccaaagagtc	agaggacttc	attgtggagc	aatataaaca	1860
tctccgccag	agagatggtt	ctggagtgac	caagtcttca	tggaggatta	cagtgcgaca	1920
gcttgagagc	atgattcgtc	tctctgaagc	tatggctcgg	atgcactgct	gtgatgaggt	1980
ccaacctaaa	catgtgaagg	aagctttccg	gttactgaat	aaatcaatca	tccgtgtgga	2040
aacacctgat	gtcaatctag	atcaagagga	agagatccag	atggaggtag	atgagggtgc	2100
tggtggcatc	aatggctcatg	ctgacagccc	tgtctctgtg	aacgggatca	atggctacaa	2160
tgaagacata	aatcaagagt	ctgctcccaa	agcctcctta	aggctgggct	tctctgagta	2220
ctgccgaatc	tctaacctta	ttgtgcttca	cctcagaaaag	gtggaagaag	aagaggacga	2280
gtcagcatta	aagaggagcg	agcttggtta	ctgggtacttg	aaggaaatcg	aatcagagat	2340
agactctgaa	gaagaactta	taaataaaaa	agaatcata	gagaaagtta	ttcatcgact	2400
cacacactat	gatcatgttc	taattgagct	caccagggct	ggattgaaag	gctccacaga	2460
gggaagttag	agctatgaag	aagatcccta	cttggtagtt	aaccctaact	acttgctcga	2520
agattgagat	agtgaagta	actgaccaga	gctgaggaac	tgtggcacag	cacctcgtag	2580
cctggagcct	ggctggagct	ctgctagggg	cagaagtgtt	tctggaagtg	atgcttccag	2640
gatttgtttt	cagaaacaag	aattgagttg	atggctcctat	gtgtcacatt	catcacaggt	2700
ttcataccaa	cacaggcttc	agcacttcct	ttgggtgtgt	tcctgtccca	gtgaagttgg	2760
aaccaaataa	tgtgtagtct	ctataaccaa	tacctttgtt	ttcatgtgta	agaaaaggcc	2820
cattactttt	aaggtagtg	ctgtcctatt	gagcaataa	ctttttttca	attgccagct	2880
actgctttta	ttcatcaaaa	taaaataact	tggtctcg			2917

<210> 413

<211> 5701

<212> DNA

<213> Homo sapiens

<300>

<308> NM_001282

<400> 413

ctgcccacca	tctttgtccc	tggaagaagt	ggttttgcgc	agtggcttag	acctagaaaa	60
gaatcgtgac	gggcaggaaa	ccattacacc	accacctggg	ctgtgctctc	cggctcccgc	120
cgccaccccc	gccctcgect	tcgctccgc	tccggtgcac	attaaagatc	caaagtcag	180
actgactcca	agtatttcac	aaccaataaa	aaaggagaaa	tatttgaact	aaaagctgaa	240
ctcaacaatg	aaaagaaa	aaagagaaa	gaggctgtga	agaaagtgat	tgctgctatg	300
accgtgggga	aggatgttag	ttctctcttt	ccagacgtag	tgaactgtat	gcagactgac	360
aatctggaac	taaagaagct	tgtgtatctc	tacttgatga	actacgcca	gagtcagcca	420
gacatggcca	tcattggctgt	aaacagcttt	gtgaaggact	gtgaagatcc	taatcctttg	480
attcgagcct	tggcagtcag	aaacctgggg	tgcatccggg	tagacaaaat	tacagaatat	540
ctctgtgagc	cgctccgcaa	gtgcttgaag	gatgaggatc	cctatgttcg	gaaaacagca	600
gcagtcctgcg	tggcaaaaact	ccatgatatc	aatgcccaaa	tgggtggaaga	tcagggattt	660
ctggattctc	tacgggatct	catagcagat	tcaaatccaa	tgggtggtggc	taatgccgta	720
gcggcattat	ctgaaatcag	tgagtctcac	ccaaacagca	acttacttga	tctgaaccca	780
cagaacatta	ataagctgct	gacagccctg	aatgaatgca	ctgaatgggg	ccagattttc	840
atcctggact	gcctgtctaa	ttacaaccct	aaagatgatc	gggaggctca	gagcatctgt	900
gagcgggtaa	ctccccggct	atcccatgcc	aactcagcag	tgggtgctttc	agcggtaaaa	960
gtcctaata	agtttctaga	attgtttacct	aaggattctg	actactacaa	tatgtctgtg	1020
aagaagttag	cccctccact	tgtcaactttg	ctgtctgggg	agccagaagt	gcagtatgtc	1080
gccctgagga	acatcaactt	aattgtccag	aaaaggcctg	aaatcttgaa	gcaggaaatc	1140
aaagtcttct	ttgtgaagta	caatgatccc	atctatgtta	aactagagaa	gttgacatc	1200
atgattcaggt	tggcatctca	agccaacatt	gtccaggttc	tggcagaact	gaaagaatat	1260
gctacaggtg	cttggtttga	ctttgttcga	aaagctgtgc	gggccatttg	acggtgtgac	1320
atcaaggttg	agcaatctgc	agagcgctgt	gtaagcacat	tgcttgatct	aatccagacc	1380
aaagtgaatt	atgtgggtcca	agaagcaatt	gttgctcatc	gggacatctt	ccgcaaatac	1440
cccaacaagt	atgaaagtat	catcgccact	ctgtgtgaga	acttagactc	gctggatgag	1500
ccagatgctc	gagcagctat	gatttggatt	gtggggagaat	atgctgaaag	aattgacaat	1560
gcagatgagt	tactagaaag	cttcctggag	ggttttcacg	atgaaagcac	ccaggtgcag	1620
ctcactctgc	ttactgccat	agtgaagctg	tttctcaaga	aaccatcaga	aacacaggag	1680
ctagtcacag	aggtcttgag	tttggcaaca	caggattctg	ataatcctga	ccttcgagac	1740
cggggctata	tttattggcg	ccttctctca	actgaccctg	ttacagctaa	agaagttagtc	1800

ttgtctgaga	agccactgat	ctctgaggag	acggacotta	ttgagccaac	tctgtctggat	1860
gagctaatact	gccacattgg	ttcttttgcc	tctgtgtatc	ataagcctcc	caatgctttt	1920
gtggaaggaa	gtcatggaat	tcatcgtaaa	cacttgccaa	ttcatcatgg	gagcactgat	1980
gcaggtgaca	gccctgttgg	cactaccact	gcaacgaacc	tggacagacc	tcaggttatc	2040
ccctctcaag	gtgatcttct	aggggatctt	ttaaaccottg	acctcgggtcc	cccagtcaat	2100
gtgccacagg	tgtcctccat	gcagatggga	gcagtggtatc	tcctaggagg	aggactagat	2160
agtctgggtg	gacaatcctt	catcccatca	tcggtgcctg	caacctttgc	tccttcacct	2220
acacctgctg	tggtcagcag	tggaactgaat	gacctgtttg	aactctccac	agggataggc	2280
atggcaactg	gtggatatgt	ggctcctaag	gctgtctggc	tacctgcagt	aaaggctaaa	2340
ggcttgagaa	tttccggaac	atttactcac	cgccaagggc	acatctatat	ggaaatgaac	2400
ttcaccaata	aagctctgca	gcacatgaca	gatttttgcaa	tccagtttaa	caaaaatagc	2460
tttggtgtca	tccccagcac	tcctctggcc	atccatacac	cactgatgcc	aaaccagagc	2520
attgatgtct	ccctgcctct	caataccttg	ggcccagtc	tgaagatgga	acctctgaat	2580
aacctccagg	tggtctgtgaa	aaacaatatc	gatgtcttct	acttcagctg	cctcatccca	2640
ctcaatgtgc	ttttttaga	agatggcaaa	atggagcgcc	aggtcttctt	tgcaacatgg	2700
aagtatattc	ccaatgaaaa	tgaacttcag	tttcagatta	aggaatgtca	tttaaagtct	2760
gacactgttt	ccagcaagtt	gcaaaacaac	aatgtttata	ctattgccaa	gaggatgtg	2820
gaagggcagg	acatgctgta	ccaatccctg	aagctcacta	atggcatttg	gattttggcc	2880
gaactacgta	tccagccagg	aaaccccaat	tacaogctgt	cactgaagtg	tagagctcct	2940
gaagtctctc	aatacatcta	tcaggtctac	gacagcattt	tgaaaaacta	acaagactgg	3000
tcagtagccc	ttcaaccatg	ctgtgatcgg	tgcaagtcaa	gaactcttaa	ctggaagaaa	3060
ttgtattgct	gcgtagaatc	tgaacacact	gaggccacct	agcaaggtag	taactagtct	3120
aacctgtgct	aacattaggg	cacaacctgt	tggaatgttt	tagcttccctg	tgaacatttg	3180
taaccactgc	ttcagtcacc	tcccacctct	tgccacctgc	tgtctctatc	tgtccttact	3240
tgtgggcttc	tccatgctgt	gccaatggct	ggctttttct	acacctctt	ttgagtgtag	3300
tttggtattt	tgtaatgtg	agctcatttc	aaaagcagaa	aaagacaaca	aatattaaag	3360
caagggaaaag	tgtaaactgaa	acactgcact	ttactgtttt	atacttttgt	acatatgaga	3420
aatcaaggga	ttagtgcac	cagtagaagg	cattgaaatg	actgtcatta	accacacagt	3480
cctggaggca	gagatgcagt	tacctacctt	agcttttgat	gggttctctt	acctgtagta	3540
gcctttatcc	tggtcatttg	gattttcagt	tgtctttttt	ctttttttcc	cctccaaact	3600
ccctttctct	ggccaagcct	tcattgttcc	ccttttccat	attataatct	catttgattg	3660
ctctgcagtt	gggaacgggtg	atcttcttga	atgatgtttc	agtgtgcaaa	aactatagag	3720
cctgtcagca	ccaaagctga	cagaagttat	acottactcc	tttcttttcc	cctgaacaaa	3780
cctgtctaata	ccactaattc	aggaatttga	gtagagatgg	ggaacaagaa	cccagatgct	3840
gtccccctac	ccccctctcct	gtattttctca	ggtccagttc	aaatctaaaa	ttctactttt	3900
agagttgaaa	cagagtaata	acttatctaa	cctctttttc	ctacaaagga	gaaagataaa	3960
aggcaaaaag	gttaccgcca	aggcccgctc	gctgtgtagt	ggcaaagccg	agaccgagtc	4020
tcctaagtcc	ccgtcagtgt	ggttttcacc	acaggactgt	ctcttgctgt	tttcccttaa	4080
tgccttctcc	tgccttttct	gtgcctagtt	tttggctctt	cacataattc	atattgattt	4140
tgcgctctg	tattattggca	tcaggtggca	gctgaatata	ttttgaatta	ctcgaaggta	4200
aagccagatg	ccagaatgaa	ggtgtagcca	gtgtttccca	tatgcccctg	gagccccact	4260
tattgaggcc	agcagaatag	gtgcagagat	gaagtgtgct	tagagatgtt	gcaaatgctc	4320
tttatccctt	cagctctctg	atctgctctt	tcttcatgat	acttagtctg	cagggcatat	4380
taagatcatc	ccagaggttc	aggcagttcc	tgtcatctct	gaaaagactg	ggggatatga	4440
aatcttcccc	ctacccctac	taatgcgttg	gatattgatt	ttcaaagaat	gcttcatgcc	4500
caaaataacca	gcctgttttag	cagtgttaca	ctgtttgatc	tgcggggact	tgttgcatgt	4560
cctggcacc	aatattcagg	gtccatgact	aagactggtc	ttctcagatg	ccctgcttaa	4620
atcaggggca	cttcaggctc	cacaggcgct	atgttggtg	gagacctaac	tcactggact	4680
cagaggagga	atcgtggaaa	acaagagcaa	aactacccca	cacccctatt	tcattgtctga	4740
aataaccctg	tttcatacca	gttgcaaaag	ttgtggggag	cggtcccaca	aagcactttc	4800
ttaaaccttg	agaattctca	agagaaaaat	atltggggaa	ggaggggagg	aatatgtccc	4860
ttgcacacca	cccctgaagc	acatggcagt	aggaacacgc	ataggattgt	atgtggggag	4920
tggaataggtc	ggtgatgtgt	ggagcggaag	agcagggttg	taaagttccc	ttcttgggac	4980
ttattcctgg	agtcagtgga	tacaagtagt	gcagaaggtt	cacactgcaa	atagtgttct	5040
catctcaaa	caaactatca	ttccagaagg	aaaagtgtgt	cagggcaagc	agacaacaca	5100
atttcctatc	agaatatgtc	cctcaacccc	cgaacaaagg	cttctctcag	cctccccacc	5160
agtgatggat	aacagctcct	attctcagct	gacctgactg	agccaaccca	tgaactcttc	5220
actccttggg	gaagccacct	cccatcacac	ccctgagcag	agttaggagg	gaattctact	5280
tcccataaaa	ggacctctcc	tgagaggcaa	aacctgttgc	ctccaccacg	gcttccctct	5340
tggtctcatc	caagcttggc	caaattgggg	aagtgggatg	gaggttgccc	tgcattcccc	5400
ctcctctgcc	tgagtgtgtc	tttgtaatgt	cagctggcat	catacaaaag	gcaggagaag	5460

```

caaacaccca gaactctttt gctggtcaga gattccctga gtgtctgtcc tcaccaagc 5520
ctgctctgtg tctgtgttgt gaagcttgag actctggaaa gaaatgggga gggggggcag 5580
gggaaatggt gccctaagaa tgcttctcat tcctctgttc ttattgggtc ctgtttttcg 5640
ggagggtggg gggtggggga agcttgacct tgtgtcttcg tcaataaact cacatttaca 5700
c                                                    5701

```

```

<210> 414
<211> 1165
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> Contig56457_RC

```

```

<400> 414
gcaagtcttt tttttttttt taataaaacc aacataacag aatgtagaaa tctaaacgaa 60
catttctccc tcaaagtagt tacatcetta gctccaacaa aactactatt gttggagact 120
tatttagaac tctgttttga ggaaaggcct taaaagggtg tttatgagcc acattagaaa 180
atacccttct aaactttggg ctgttttaaaa acagaaaatc catactcaga ggatgacacc 240
gagaaaatta aaaagggtgt gcagctctca cggagggttc ctaaagacct ctagcaagtt 300
ctgagcaaca gcatcaccat tggagaaacg tgtagcctcc caggacagcc actttccatg 360
ggacggcatt catttgata gaaaagttct gctgtctgtt taaaaagaa aaagaaaaaa 420
aaaaaatcag ccacactcct ttatagtcac acttgcattt gattaaaaaa aaaaacaaca 480
acaacaacac tcacacgtgg ttcttgtatg aattctaatt aagagacatt atattcagaa 540
ttatagcact catagtata attttcagaa gactataaca gcttctttgc cacttaaggg 600
gtacagaggg tgcagctgtt tacctgaatg aattcttaaa taagtacta ctaccactct 660
aattttccta gttgtagctc tgttacaggt tactgaccgc caagtccatg ttccagacat 720
tctttgcaa cacagattaa ttacaagaga gccaccccca aagtcttagg gcgaattggg 780
aaagggcgcg ttaaattgag aaacaggcca ttcccttgga aaagtgtgta gttttctttc 840
ccaatacatc ttctttttat tgagaacttc cttttttatc ctgtgttctc cctctgacct 900
ctcacactct cgctctatct gctggatttt ggccacatgc acctcattca tttgtatcag 960
ttgcagttgt aagatggaca tttgttgatg gtgttcttcc tcatgcatct tttttaaagc 1020
accttctga gaggttttgc tcctgtagtt tttattggct gttattctga cagctgaaca 1080
cgagggttca gggtgagagg tccctcaca tactggaaaa tgtatgaact ctttctcatc 1140
atcgcacagt tctctatttg aaaca                                1165

```

```

<210> 415
<211> 867
<212> DNA
<213> Homo sapiens

```

```

<300>
<308> NM_020386

```

```

<400> 415
cgtttcagcg tggcgggcgt ggtgctggcg ttggccctgg aggacggccc cgagtgatgg 60
ctggcgectg cctcccgggt gtctcccggt tacagatgga gtcgtcccg cgcgcggcgc 120
ggcaaggctc gcagctgcga ggccaagaga gaccccagga cacacacagc tgcctcccgg 180
tgcgagaaga agaccccggc ttgagagtga gatggcgttt aatgattgct tcagtttgaa 240
ctaccctggc aaccctgcc caggggactt gatcgaagtg ttccgtcctg gctatcagca 300
ctgggcccgt tacttgggtg atggttacgt tatcaacata gcacctgtag atggcattcc 360
tgcgtccttt acaagcgcca agtctgtatt cagcagtaag gccctggtga aaatgcagct 420
cttgaaggat gttgtgggaa atgacacata cagaataaac aataaatacg atgaaacgta 480
ccccctctc cctgtggaag aaatcataaa gcggtcagag tttgtaattg gacaggaggt 540
ggcctataac ttacttgtca acaactgtga acattttgtg acattgcttc gctatggaga 600
aggagtttca gagcaggcca accgagcgat aagtaccgtt gagtttgtga cagctgctgt 660
tgggtgttcc tcatctctgg gcttgtttcc aaaaggacaa agagcaaaat actattaaca 720
atttaccaa gagatattga tattgaagga atttgggagg aggaaaagaa acctgggggtg 780
aatacttatt ttcatgtcat cattactgtt ccagattcct atgatggatg gcagactctt 840

```

taataaattg cttactgata ttatctt

867

<210> 416

<211> 3432

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> 2741

<223> n = A,T,C or G

<300>

<308> NM_014889

<400> 416

```
ctcgtcatgc gcaatgtggc gctgcgccgg gcggcagggc ctgtgtgtgc tgaggcggct 60
gagcggcggg catgcacacc acagagcgtg gcgatggaac agtaaccggg cttgtgagag 120
ggctctgcag tataaactag gagacaagat ccatggattc accgtaaacc aggtgacatc 180
tggtcccgag ctgttcctga ctgcagtga gctcaccat gatgacacag gagccaggta 240
tttacacctg gccagagaag acacgaataa tctgttcagc gtgcagttcc gtaccactcc 300
catggacagt actggtgttc ctacacattct tgagcatacc gtcctttgtg ggtctcaaaa 360
atatccgtgc agaaaccctt tcttcaaaat gttgaaccgg tccctctcca cgttcatgaa 420
cgccttcaca gctagtgtt atactctgta tccattttcc acacaaaatc ccaaggactt 480
tcagaatctc ctctcgggtg atttggatgc caccttttcc ccatgtttac gcgagctgga 540
tttctggcag gaaggatggc ggctggaaca tgagaatccg agcgaccccc agacgccctt 600
ggtctttaaa ggagtcgtct ttaatgagat gaagggagcg ttacagaca atgagaggat 660
attctcccag cacttcaga acagacttct tcccgaccac acgtactcag tggctctccg 720
gggtgaccca ctgtgcatcc cggagcttac atgggagcag cttaaagcagt ttcatgccac 780
tcactatcac ccaagcaatg ctaggttctt cacgtacggt aattttccat tagaacagca 840
tctgaaacaa attcacgagg aagcactgag caaattccag aaaattgaac caagcaccgt 900
ggtgccagct cagacaccct gggacaagcc tagggaattc cagataacat gtggcccgga 960
ttcatttgct acagatccct ctaaacaaac aaccgtcagc gttagcttcc tcttaccgga 1020
catcacccac acatttgaag ccttcacatt aagtcttctg tcttactctt tgacttctgg 1080
gcccaattct cctttttaca aagccttgat tgaatctggc cttggcacag aattttctcc 1140
tgatgttgga tataatggct acacgagggg ggctactttt agtgtcggcc tccaagggat 1200
tgtggagaaa gacattgaga ccgtcagaag cctcatagac agaacgattg atgaagtagt 1260
tgagacaagg attgaagatg atcgaattga ggctttactt cataaaattg aaatacagat 1320
gaaacatcag tctaccagct ttgggctgat gctgacatca tacatagctt cttgctggaa 1380
ccatgatggg gacctgtgg agctcttgaa gttgggaaat cagttagcta aattcagaca 1440
gtgcctgcag gaaaatccaa aatttttgca agaaaaagta aaacagtatt ttaagaataa 1500
ccagcataag ctgactttat cgatgaggcc agatgacaag tatcacgaga agcaggcaca 1560
ggtggaagcc acgaagctca agcagaaggt cgaggctctg tcccccgag acaggcagca 1620
gatctacgag aaaggtctag aattacggag tcaacaaagc aaacctcaag atgcctcttg 1680
tctgccagcg ttgaaagtgt ccgatatgta acccaccata cctgtcacag agttggacgt 1740
ggtcctgaca gctggagata tccctgttca gtactgcgcc cagcccacca atggcatggt 1800
gtatttccgg gccttctcca gcctgaacac actccccgag gagctgaggc cctatgtgcc 1860
cctcttctgc agcatcctca ccaagctggg ctgcggcctt cttgactacc gggagcaggc 1920
tcagcagata gaattgaaga ccggagggat gagtgtctct cccacgtgc tccccgacga 1980
ctcacacatg gacacctacg agcaggtagg tgtgttttct tctctctctt gcctggatcg 2040
aaacctgcca gacatgatgc agctatggag tgaaatattt aacaaccgtg gctttgaaga 2100
agaggagcac ttcaaggtgc tggatgaagat gaccgcccag gagctcgcca atggaattcc 2160
tgactctggg ccctgtacg catccatcag ggcaggccgg accctcacg ccgcagggga 2220
cctgcaggag accttcagcg ggatggatca ggtgcggctg atgaagagga ttgcagaaat 2280
gacagatata aaacctatcc tgaggaaagt cccgcgtatc aagaaacact tgttaaattg 2340
tgataatatg aggtgttcag tgaatgcgac tcctcagcag atgcctcaga cagaaaaagc 2400
ggtcgaagac ttcttagaaa gcatcggtcg gagtaaaaag gaacggaggc ctgtgcgccc 2460
acacacgggc gagaaacctg tgcccagcag ctctggtgga gatgcccacg tccccatgg 2520
ctcccaggct attaggaagc tggatcatgga acccaccttc aagccctggc agatgaagac 2580
tacttctctg atgcccttcc cggatgaatta tgcattccga ctgtccccta 2640
cacggaacca gatcatgcca gtcttataat ccttgacagt ttgatgactg ccaattctt 2700
```

gcatacagaa attcgagaaa aaggcgggtgc ttatgggtgga ngcgcaaaac tcagccacaa 2760

tgggattttc accctttact cttacaggga cccaaatata atagagacgc tccagtcttt 2820
 tgggaaggct gtgcactggg ctaagtctgg aaaattcaca cagcaagaca tcgacgaagc 2880
 caaactttct gtcttctcaa ccgtagatgc tctgtcgcct ccttcagaca aaggaatgga 2940
 ccacttcttg tacggcctct cggatgagat gaagcaggcc cacagagagc agctctttgc 3000
 tgtcagccac gacaagctcc tggcgtgag cgataggtag ctcggcactg ggaagagcac 3060
 acacggcctg gccatcctcg gaccgagaa cccgaaaatt gccaaaggac catcctggat 3120
 catccgatga gcagccgtgg cgctcgactg cacaggagcc cgagacaata cacctccaag 3180
 ctgaatatga aaagtcagaa atgctactgc tttttccaag aatattatgt cattgagtgt 3240
 cgccaaagcc cttgactggc gagtcaaaaa ctcagatcta tcttaagagt gaccaggaag 3300
 aggttcattg aaataatcat gcatgaagcg ccaaagatgc accatgtaga attttcaactt 3360
 tgtactggca ggctcgtttt acctgattct agaattattt agaattctaa aataaagggc 3420
 aactctgact ta 3432

<210> 417

<211> 817

<212> DNA

<213> Homo sapiens

<300>

<308> NM_006681

<400> 417

agtctgcgt ccggggcccc aggcgcagca gggcaccagg tggagcacca gctacgcgtg 60
 gcgcagcgca gcgtccctag caccgagcct cccgcagccg ccgagatgct gcgaacagag 120
 agctgccgcc ccaggctgcc cgccggacag gtggcccgcg cgtccccgct cctgctgctg 180
 ctgctgctgc tcgcctgggt cgccggcgcc tggcggaggt ctccaatatt acctcaagga 240
 ttacagcctg aacaacagct acagtgtgtg aatgagatag atgatacttg ttcgtctttt 300
 ctgttcattg attctcagcc tcaggcatcc aacgcactgg aggagctttg ctttatgatt 360
 atgggaatgc taccaaaagg tcaggaacaa gatgaaaaag ataatactaa aagggtctta 420
 tttcattatt cgaagacaca gaagttgggc aagtcaaatg ttgtgtcgtc agttgtgcat 480
 ccgttgctgc agctcgttcc tcacctgcat gagagaagaa tgaagagatt cagagtggac 540
 gaagaattcc aaagtcctct tgcaagtcaa agtcgaggat attttttatt caggccacgg 600
 aatggaagaa ggtcagcagg gttcatttaa aatggatgcc agctaatttt ccacagagca 660
 atgctatgga atacaaaatg tactgacatt ttgttttctt ctgaaaaaaa tccttgctaa 720
 atgtactctg ttgaaaatcc ctgtgttgct aatgttctca gttgtaacaa tgttgtaaat 780
 gttcaatttg ttgaaaatta aaaaatctaa aaataaa 817

<210> 418

<211> 5210

<212> DNA

<213> Homo sapiens

<300>

<308> NM_007203

<400> 418

ccagaggctg agcccggggc agcccgcctg gcgcacagct ctgcccgccg cccctggagc 60
 ggatccccag ccgtccctgg gcgttctccc ggggtctagc aaagccccc ggggtgcca 120
 tcagtttctt tgggtgacta cagcgtgtgt ttttctttc ctctttccc tgctgtgtg 180
 ccttcttcca ggatggcaga ggcggaattg cacaaggaaa ggctgcaagc catagcagaa 240
 aaaagaaaga ggcagactga aatagaaggc aagcgacaac agcttgacga gcagatactt 300
 ctgctgcagc attccaagtc caaagtgtct cgggagaaat ggctgctgca gggcataccc 360
 gctggaactg ccgaagagga ggaagccagg aggcggcagt ctgaagagga tgagttcaga 420
 gtcaagcaac ttgaagataa cattcagagg ctggagcaag aaatacaaac gctagaaagt 480
 gaagagtccc agatatctgc caaagagcaa atcatcctag agaaactgaa ggaaacagaa 540
 aaatccttca aggactttca gaagggttcc tcagtagcgg atggagatgc agtaaattac 600
 atttctctcc agcttcccga cctgccaatc ctctgttcac gaacagcaga accatcacct 660
 gggcaggacg ggaccagcag agcggctgga gtcgggtggg agaattgtct gctaaaggaa 720
 ggtgagtcag cctcgaacgc cacagaaaca tccggcccag acatgactat caagaagcct 780

```

ccccagcttt ctgaggatga tatctggcta aaaagcgagg gagacaacta tagtgccacc 840
ctcctggagc ctgctgccag ctctcttttc ccagatcaca aaaacatgga aattgaggtg 900
tctgttgagc aatgtaaaag tggttcctga atcacctcta ccccatatcc catggaccat 960
ccctccgctt tctattcacc cccgcataat ggctctctta ctgatcacca cgaatccctg 1020
gataatgatg ttgccagaga gatccgctat ctagatgagg tgctagaggc caactgctgt 1080
gattctgctg tggatggaac gtacaatgga acatcctccc cagagcctgg tgcagtgggt 1140
ctggtgggag gcctaagccc ccctgtccac gagcgacccc agccagaacc cactgaaaga 1200
acagctagcc ggagggcacc tcttcacatc gagctcagta atagcagccc tgaccccatg 1260
gcagaggcag aaagaacaaa tggccattcc ccagccagc ctagagatgc gctggggagc 1320
agcctgcagg tgctgtcag cccagctcc acaaccagct cacggtgttc ttcccagat 1380
ggagagttca ctctcaccac actgaaaaag gaggccaagt ttgagctgag tgccttccat 1440
gaggacaaga agccctccaa gctctttgag gatgacgagc atgagaaaga acaatactgc 1500
attagaaaag tgaggccttc agaggagatg ctggagctgg aaaaggagag gagagagctc 1560
atccgcagcc aggcctgcaa gaagaatcct ggcattgcag caaatgggtg gaatcccccg 1620
caggaaaaaa ccatcgagga gcagctggac gaggaacatc tggagtgcga caaaaagtac 1680
aaggagcgca agcagcagag ggacaggtgc gaacagttgc tgctgcagaa gcagttacag 1740
cagcagcagc agcagcccc atcgcagctc tgcacagccc ctgcctcttc tcatgaacgc 1800
gcaagcatga ttgacaaagc aaaggaggac attgtcacag agcagataga tttctctgct 1860
gctcgcaaac aatttcagct gatggagaat tccaggcaag cggtggccaa gggccagagt 1920
acaccagagc tgttctccat caagcctttc tacaggcctc tggggtcagt caactcagac 1980
aagccactga ctaatccgag accacottct gtccggggagc ctccagaaga cagtgtgtgc 2040
tcagccgcca agggacagaa atcccccggt gccctggaga ccccatcggc agcaggaaagc 2100
cagggaacaa cagcctctca ggggaaggaa gggccctaca gcgagccttc taaacgtggg 2160
cccttatcta aactgtgggc taaggatgga gaatttacga gcgcccggc tgtcctcact 2220
gtggtcaagg atgatgacca tgggattttg gatcagttct caagatctgt caatgtctcc 2280
ttgacccaag aggagcttga ctctggtctg gacgaattgt cggtgaggtc tcaggatacc 2340
acagtcctgg agacctatc caatgatttc agcatggaca acatcagtga cagcggggca 2400
tccaatgaga caaccaatgc cctccaggaa aattcactgg ctgatttttc tctgcccag 2460
acaccacaaa ctgacaaccc ctccagagggc cgaggagaag gcgtctccaa gtcatttagt 2520
gatcatggtt tctattcccc ttcctccacg ctgggggact ctccgttggg tgatgacccc 2580
ttggagtatc aggcctggct cctggtgcag aatgccattc aacaagccat agccgagcag 2640
gtggataaag ctgtgtccaa aaccagcagg gatggagcag agcaacaggg acctgaagcg 2700
actgtagagg aagctgaagc tgcggctttc ggctcagaaa agcctcagag catgtttgag 2760
ccacctcagg tgtctctctc tgttcaagag aaaagggatg tattacaaa gattctgcct 2820
gctgaagaca gggcgctcag ggaaggggg ccccccagc cactgccagc tgtgcagccc 2880
agtggcccga ttaacatgga ggagaccagg cccgaaggaa gctatttcag caagtactcg 2940
gaggcagctg agctgagaag cacagcctcc ctccctggca ctcaagaatc tgacgtgatg 3000
gttgggcctt tcaagctgag gtccagaaa cagcggactt tgtccatgat tgaggaaag 3060
atccgagcag ctccaggaaa ggaagaggag ctgaagaggc agagacaagt cttgcagagt 3120
acgcagagcc ccaggacaaa gaatgcccc tccactgcct ccagaacatg ctacaaaact 3180
gctccaggga aaatagagaa agtcaaacct cctccatccc ccaccactga aggccccagc 3240
ctgcagcctg acttagcccc tgaagaggct gccggaaccc agcggcccaa gaatctgatg 3300
cagaccctca tggaagacta tgagacacac aaatctaaaa ggcgcgagag aatggatgat 3360
agtagttaca ctctcaagtt actgtcttgc aaggtgactt ccgaggtcct cgaggccaca 3420
cgggttaatc gaagaaagag cgcactggct ttgcgctggg aagcagggat ctatgccaac 3480
caggaggaag aagacaacga ataaacttcc ttcaaccag gaagcgtctt tgggtgcttg 3540
gagaccaaga aaccaagaaa ttaacaactg aaagcatttt aatggactat ttattaaagt 3600
gcaaacaaac tcagcaatcc ttatgtagac cagaagctgc aatatacaat gatgaaaatg 3660
aaacgaaaag gaagtcccc actcaatccc agtgtctggt ttggggccaa tctacaatgg atccaaaagg 3780
acaaaacatg atcaacatag ttgctgcagg cggaactgag accagcctct 3840
ctgggcgatg aaggacaatg gtggtagggc cacttagaga gcagtaggc ccaggctggt 3900
gcagcctaca tggcacagga tatttccctg caccagtttc ccacctgtca cttgcagcat 3960
cagagggtgt tgcattaact ccattttctc cagcagccat aaaatgaatg gcgtttctct 4020
ccatatttgc ccagaagagg cagtatgat tccacttctc ctacatttaa ttatttttaa 4080
ttgagaacaa ctcatgatca caattcataa aatccatatt ttaagacaat agttcttcat 4140
gtcacaaaac caatctattc tacatgagga ggctgagaaa ggggaagaga aaaacaaacc 4200
aaatggatga tgttagcttt aggatcgtgt ctgcttataa tattcactga tctgcagttt 4260
atgagaacac attttcacag ttatgtttct tcataggaaa actatattta gtgggcaaca 4320
actattttta tgatgggatg ggggagtata tacacgtata gaatctgtac gcgttgaaca 4380
acttggttca agatggtggg ggcattttta gagcggcaat aattgaaaaa aaaggcgaac 4440

```

```

tctgccttg agaggttagat gataagaaat aaaaagggtgt ttataactat tttgtattat 4500
aaagtgggcc ttagagatag gaagaagaat gatggattcc ttttggatca atcagaaagg 4560
aaacacgaaa gaaaagtcag gaaggttagag agagaaaaag ggagggaagg agaaagaatg 4620
ggaataaaat aaggaggtta gagatactat ttttgctgag caaccagtgt gtttcaggat 4680
gatacaaaaga aaaatataga atagaaataa gtgcaggctt ggaatcagct acaaatccta 4740
aagatgggggt gtgtgtggat gtgtgtgtgt gtgtgtgtac accattgtgt gtttgtaaaa 4800
tgtgtatgtt catgagtaag ggtgtgtgtg tgtgtgtatt aaaattccag agtgaccgtg 4860
gcacttgggt gtacaggtaa ttctccaga gctgtttgct ggcttcagga gtggagttag 4920
aatttccttt ttatgaaaag ggatataaag gcaccgagct gatgcagtat ttgtaatat 4980
aagttgacct aacaaggat ttgcatgagt cacaattaca aagttttgag cggttttgta 5040
atttgacatt taggaaagtc tctatattt tctcatactt tacattcatg cttagtatac 5100
tatagaggat gccagcttta atctttctgt catttaaagc aatatgataa gggatttcaa 5160
taattgggtg ccctaaattt ctggatgaga aaattttcaa ttctggccag 5210

```

<210> 419

<211> 449

<212> DNA

<213> Homo sapiens

<300>

<308> Contig63102_RC

<400> 419

```

ctatcccatt ttttcattta tttttatgag gagcctcaaa aaatagaagt ggccttggtg 60
gggaaggaat gtctgtatgc tgagtacagt atggcaaatt cttttgcctc agtttcccta 120
actccatcca gtgcatatga aaatgtctat gtgcgtacaa ggtaggtctg gactaagctc 180
tgccctccg gttgttttcc cattgcggga gcctaagcca ggggtgccag gactcctgtg 240
tggtgtgtg gcctgatgcc cacagctgat gattcaggaa ggaggggctt ggagtctgct 300
gcagactctc ccgcccccta cagcccaaac cgggcactgc agcaatgagg tggtcagctc 360
agggagaggt ccgacaagtt ctcggcacaa tgctgcagga agtcaaagtc aggcacggag 420
aagggcacgc gggaccactt tttggcctg 449

```

<210> 420

<211> 3044

<212> DNA

<213> Homo sapiens

<300>

<308> NM_003981

<400> 420

```

gtttgcgggt ggttgttgct ctcggggccc tgtggagtag gtctggacct ggactcacgg 60
ctgcttgagg cgtccgccat gaggagaagt gaggtgctgg cggaggagtc catagtatgt 120
ctgcagaaag ccctaaatca cttcgggaa atatgggagc taattgggat tccagaggac 180
cagcggttac aaagaactga ggtggtaaag aagcatatca aggaactcct ggatatgatg 240
attgctgaag aggaaagcct gaaggaaaga ctcatcaaaa gcatatccgt ctgtcagaaa 300
gagctgaaca ctctgtgcag cgagttacat gttgagccat ttcaggaaga aggagagacg 360
accatcttgc aactagaaaa agatttgccg acccaagtgg aattgatgagc aaaacagaaa 420
aaggagagaa aacaggaact gaagctactt caagagcaag atcaagaact gtgcgaaatt 480
ctttgtatgc ccactatga tattgacagt gcctcagtgc ccagcttaga agagctgaac 540
cagttcaggc aacatgtgac aactttgagg gaaacaaaagg cttctaggcg tgaggagt 600
gtcagtataa agagacagat catactgtgt atggaagaat tagaccacac ccagacaca 660
agctttgaaa gagatgtggt gtgtgaagac gaagatgcct tttgtttgtc tttggagaat 720
attgcaacac taaaaagtt gctacggcag ctggaaatgc agaaatcaca aaatgaagca 780
gtgtgtgagg ggtcgcgtac tcaaatccga gagctctggg acaggttgca aatacctgaa 840
gaagaaagag aagctgtggc caccattatg tctgggtcaa aggccaaagg ccggaaagcg 900
ctgcaattag aagtggatcg gttggaagaa gtgaaatgc aaaacatgaa gaaagtgaat 960
gaggcaattc gagtggagct gtttcagtac tgggaccagt gcttttatag ccaggacgag 1020
agacaagctt ttgccccttt ctgtgtctgag gactacacag aaagtctgct ccagctccac 1080

```

```

gatgctgaga ttgtgcggtt aaaaaactac tatgaagttc acaaggaact ctttgaaggt 1140
gtccagaagt gggaagaaac ctggaggcct ttcttagagt ttgagagaaa agcttcagat 1200
ccaaatcgat ttacaaaccg aggaggaaat cttctaaaag aagaaaaaca acgagccaag 1260
ctccagaaaa tgctgcccaa gctggaagaa gaggttgaagg cacgaattga attgtgggaa 1320
caggaacatt caaaggcatt tatggtgaat gggcagaaat tcatggagta tgtggcagaa 1380
caatgggaga tgcacgcatt ggagaaagag agagccaagc aggaaagaca actgaagaac 1440
aaaaaacaga cagagacaga gatgctgtat ggcagcgctc ctogaacacc tagcaagcgg 1500
cgaggactgg ctcccaatac accgggcaaa gcacgtaagc tgaacactac caccatgtcc 1560
aatgctacgg ccaatagtag cattcggcct atctttggag ggacagtcta ccaactcccc 1620
gtgtctcgac ttctctcttc tggcagcaag ccagtcgctg ctccacctg ttcagggaa 1680
aaaacacccc gtactggcag gcatggagcc aacaaggaga acctggagct caacggcagc 1740
atcctgagtg gtgggtaccc tggctcggcc cccctccagc gcaacttcag cattaattct 1800
gttgccagca cctattctga gtttgcaag gatccgtccc tctctgacag ttccactgtt 1860
gggcttcagc gagaactttc aaaggcttcc aaatctgat ctacttctgg aatcctcaat 1920
tcaaccaaca tccagtcctg agaagccctg atcagtcaac cagctgtggc ttctgtgccc 1980
tagactggac ctaattatat gggggtgact ttagtttttc ttcagcttag gagtgttga 2040
aaccttggcc aggttccatg accatgggccc taacttaaag atgtgaatga gtgttacagt 2100
tgaaagccca tcataggttt agtggtccta ggagacttgg ttttgactta tatacatgaa 2160
aagtttatgg caagaagtgc aaattttagc atatggggcc tgacttctct accacataat 2220
tctacttgct gaagcatgat caaagcttgt tttatttcac cactgtagga aaatgattga 2280
ctatgcccac ccctgggggt aattttggca tgtatacctg taactagtaa ttaacatott 2340
ttttgtttag gcatgttcaa ttaatgctgt agctatcata gctttgtctt tacctgaagc 2400
cttgtcccca ccacacagga cagccttccct cctgaagaga atgtctttgt gtgtccgaag 2460
ttgagatggc ctgccctact ggcaagagg tgacaggaag gctgggagca gctttgttaa 2520
attgtgttca gttctgttac acagtgcat tgcctttgtt gggggtatgc atgtatgaac 2580
acacatgctt gtccgaacgc tttctcggcg tttgtccctt ggctctcatc tccccattc 2640
ctgtgcctac tttgcctgag ttcttctacc cccgcagttg ccagccagat tgggagctct 2700
tttgttccaa tggggttgagc tgtctttgtc gtggagatct ggaactttgc acatgtcact 2760
actggggagg tgttctgtct ctagcttcca cgatgaggcg ccctctttac ctatcctctc 2820
aatcactact cttcttgaag cactattatt tattcttccg ctgtctgcct gcagcagtac 2880
tactgtcaac atagtgtaaa tggttctcaa aagcttacca gtgtggactt ggtgttagcg 2940
acgtgtttta cctcatacag tacgtgtcct gtttttaaaa tatacaatta ttcttaaaaa 3000
taaattaaaa tcgtatactt acatttcaaa aaaaaaaaaa aaaa 3044

```

<210> 421

<211> 407

<212> DNA

<213> Homo sapiens

<300>

<308> Contig20217

<400> 421

```

agtgttacag acagatattg tgaaattttt ccttgggctt ctttcaattt cgttcaattc 60
tcgtccagta ttctcataac atttggatca ggtactttat cacctttagt tgtcactgtt 120
cctttctgta tcctctcgaa gcctcctgtt actcgttttc taattctctt ccagccact 180
gcacgatcta gtgtcgaatg gatgaaacag ctgagcagaa cagttcttta gctgtcggcc 240
taaagacca acttcatgac ctctcaaggc ctgcctccc ggctgcacgc agctcagctt 300
ccaacattcg ggaagggcgt gccggtatcg gaggggccc tgggaccgag tccggacgtc 360
gcgacgcgca gggggctgct gggccggcgg tctcgggaaa agtgagg 407

```

<210> 422

<211> 1170

<212> DNA

<213> Homo sapiens

<300>

<308> Contig2399_RC

<400> 422

```

aatgtttcaa caacatttat acatcttttt aattagatta gctttacaag caacttgaga 60

```

```

gctctttcat aatgaaacac tgcttttcaa attacatgac tttatagtct atctgtatat 120
gaatcctgaa aatcctttgcc tactcttgat aaaggctctgt gttttacagc tggttaatgt 180
gttgaataaa aacagcatta tgaagcttta tctttaacca ggattaaaca catataggcc 240
acaaacagtt ttaaattttg tttgaaatag agtctctgctt ggtgacaagt taatagtccc 300
actaacatga aatgaatgga tattaataaa atggccatcc tgatttcttg atcttttcac 360
aatgattcct aagagagtat ataaactttt ccaaaaaata tacaatttaa gacctgtctg 420
gcaagcaciaa cccacagata gtagaatcta atacttttta caactgtaga ggtacacatt 480
tgaatgacaa caggggctcc caaatttaaaa ggctgtggct gtacattttg acaatcaaac 540
aaacctccca tcaacataaa acaattttcta ctaaccctaa aaaagataat gcttttgagt 600
tgtggcctaa cgcgcgcac tgtactatag accagatggc aaagacaact tgcgctcctg 660
ggttgggctg gcagaggcac ttactgtaga gatgtgatgt tatcatctgg aaggaagggg 720
aggaagtaga acaaaactga atatgggcga ggaaaaaaca ataaactatc actttctcct 780
gcaaatcttc ctaagtgaga tcaactatact gaacattcaa tcccatccta agtccactgt 840
tggtcttact acacaatttt ttttccatgt tttacatgaa tagcaatggc aatcttttaa 900
taaaaattac tgagcttccc atagaaaagg tctcaaattg aatacaaact atacagatgc 960
taaaaatgtc ttaagttgaa atatacagaa aatttctgta tactcacatt ttggcaaaga 1020
agatcaagga cagagataga aaacctaaagt ggccaaggct ttacacccat aattcacgca 1080
aaatatggat accacccttt aggttttgct aaaaagccta ggaaagtga gtctaagaca 1140
tccgtcccaa ttcattcttt cctcgtgccg

```

<210> 423
 <211> 2019
 <212> DNA
 <213> Homo sapiens

<300>
 <308> NM_007036

```

<400> 423
tggtctcagc ctgcttccca ccagcaaaga ccacgactgg agagccgagc cggaggcagc 60
tgggaacacat gaagagcgtc ttgctgctga ccacgctcct cgtgcctgca cacctggtgg 120
ccgcctggag caataattat gcggtggact gccctcaaca ctgtgacagc agtgagtgc 180
aaagcagccc gcgctgcaag aggacagtgc tcgacgactg tggctgctgc cgagtgtgcg 240
ctgcagggcg gggagaaact tgctaccgca cagtctcagg catggatggc atgaagtgtg 300
gcccggggct gaggtgtcag ccttctaagt gggaggatoc ttttggtgaa gagtttggtg 360
tctgcaaaga ctgtccctac ggcaccttcg ggatggattg cagagagacc tgcaactgcc 420
agtcaggcat ctgtgacagg gggacgggaa atcgcctgaa attccccttc ttccaatatt 480
cagtaaccac gtcttccaac agatttggtt cctcaccgga gcatgacatg gcatctggag 540
atggcaatat tgtgagagaa gaagtgtgta aagagaatgc tgccgggtct cccgtaatga 600
ggaaatggtt aaatccacgc tgatcccggc tgtgatttct gagagaaggc tctattttctg 660
tgattgttca acacacagcc aacattttag gaactttcta gatatagcat aagtacatgt 720
aatttttgaa gatccaaatt gtgatgcatt gtggatccag aaaacaaaaa gtaggatact 780
tacaatccat aacatccata tgactgaaca cttgtatgtg tttgttaaat attcgaatgc 840
atgtagattt gttaaatgtg tgtgtatagt aacactgaag aactaaaaat gcaatttagg 900
taatcttaca tggagacagg tcaaccaaaag agggagctag gcaaagctga agaccgcagt 960
gagtcaaatt agttctttga ctttgatgta cattaatgtt gggatatgga atgaagactt 1020
aagagcagga gaagatgggg aggggtggg agtgggaaat aaaatattta gcccttcctt 1080
ggtaggtagc ttctctagaa ttttaattgtg cttttttttt ttttttggct ttgggaaaag 1140
tcaaaaataaa acaaccagaa aaccctgaa ggaagtaaga tgtttgaagc ttatggaaat 1200
ttgagtaaca aacagctttg aactgagagc aatttcaaaa ggctgctgat gtagttcccc 1260
ggttacctgt atctgaagga cggttctggg gcataggaaa cacatacact tccataaata 1320
gctttaacgt atgccacctc agagataaat ctaagaagta ttttaccac tgggtggttg 1380
tgtgtgtatg aaggtaaata tttatatatt tttataaata aatgtgttag tgcaagtc 1440
cttccctacc catatttatc atcctcttga ggaaagaaat ctagtattat ttgttgaaaa 1500
tggttagaat aaaaacctat gactctataa ggttttcaaa catctgaggc atgataaatt 1560
tattatccat aattatagga gtcactctgg atttcaaaaa atgtcaaaaa atgagcaaca 1620
gagggacctt atttaaacat aagtgcgtgt acttcgggtg attttcaatt taaggatatg 1680
aaataagttt ttaggaggtt tgtaaaagaa gaatcaattt tcagcagaaa acatgtcaac 1740
tttaaaatat aggtggaatt aggtatatat ttgaaagaat cttagcacia acaggactgt 1800
tgtactagat gttcttagga aatatctcag aagtatttta tttgaagtga agaacttatt 1860
taagaattat ttcagtattt acctgtattt tattcttgaa gttggccaac agagtgtgta 1920

```


atgtgtgtgg aaggcctttg aatgtaaagc tgcataagct gttagggtttt gttttaaaag 1980
gacatgttta ttattgttca ataaaaaaga acaagatac 2019

<210> 424

<211> 1720

<212> DNA

<213> Homo sapiens

<300>

<308> NM_018354

<400> 424

gactgcagat	gaaattagta	actggtggg	tcgtgggggtg	tgaatgggtgg	gcgggagcag	60
ctatgtcagt	tggtgtgttt	ctgcttatgt	tagggtaatt	gggcacggcc	tttgtgtaac	120
tggtgaatat	ctctgaacct	gggcatgaaa	cagagagatg	tcctaactct	gggtgagagg	180
aatcctcatt	tttctctgcc	ctctcactgt	ggcatcctaa	gaaaaaagtt	ttgggttcct	240
gcagcatgaa	ggagagctct	gtccccagaa	tttgggagct	ccagatttct	tccagggtgt	300
ggaggcatca	atatatcagt	ctgggaaagg	ggttcctggg	ccactccagg	agctgagttg	360
ggtggaaggt	gctgagagtg	tgggtggggg	ccacttctga	gcacccatgt	ggcaccact	420
gctggtccct	gtttgtggct	gggcactcag	gaaaatgttt	ttggtgctaa	gagtaaaaag	480
ccaaccaaca	aacacatctc	ttttttctgt	ctattcactg	gaaagtaaaa	gcagtctggg	540
cgcaggctgg	ggaccagat	ggaattcaaa	cttatgcctg	ctctcaagg	gctcacgggt	600
gctgataaac	agctggataa	aatgaagagt	ctatgagtga	gggatgcaga	gccagggaag	660
gctggtggag	tgatgccacc	agcacagggg	tatgagtttg	cagctgccaa	ggggccaagg	720
gatgagctgg	ggcctcctt	cccaatggca	tctccccctg	gtctggaact	gaagacactg	780
agcaatggtc	cccaagcccc	aaggagatca	gctccccctg	gcccagtggc	cccaaccagg	840
gagggtgtgg	agaatgcctg	cttctcctca	gaggagcatg	agacccatct	ccagaaccct	900
gggaacacga	gactgggcag	ctcaccaggt	ccccctgggg	gtgtctcctc	actgccccga	960
tcccgcggg	atgatctgtc	ccttcattca	gaggaggggc	cagccctgga	gcccgtgagc	1020
cgccccgtgg	attatggctt	tgtttccgcc	ctcgttttcc	tggtgagtgg	gattcttctg	1080
gtggtgacag	catacgccat	cccccgtag	gctcgagtca	atccggacac	agtgcagcgc	1140
cgggagatgg	aacgactgga	gatgtactac	gcccgcctag	gctccacact	ggacagggtgc	1200
atcatcgag	gcctcgggct	gctcacgggt	ggcggcatgc	tcttgctggg	gctgctcatg	1260
gtctccctgt	gcaaggcgga	gctgtaccgc	cggaggacct	tcgtccccgg	caagggtctc	1320
aggaagacct	acggctccat	taacctgcgc	atgagacagc	tcaatgggga	tgggggccag	1380
gccctggtgg	agaatgaagt	tgtccaggtc	tcagagacta	gccacaccct	ccagagggtct	1440
taagaactag	cccaccttat	ctggctgctt	tagctccagt	gctacaaggt	ccacccccctg	1500
ctcccgccca	cctgaccctt	gccaaggccc	tggggtttta	aactgagctc	acataggggc	1560
ttgtggaaga	agtactgggt	gctggaggga	gagctcgggg	cccagcccat	gccccacacg	1620
ggcaagcagc	ccactgatct	gttttgtagc	tgagggtttt	catacgggtt	tgtttgaggg	1680
atggcttctg	ctgctaaaaa	tacaaaagtt	tggaaaccgc			1720

<210> 425

<211> 987

<212> DNA

<213> Homo sapiens

<300>

<308> NM_000125

<400> 425

cggagagggg	gagaacagac	aacggggcgg	ggggagcagc	atggagccgg	cggcggggag	60
cagcatggag	ccttcggctg	actggctggc	cacggccgcg	gcccggggct	gggtagagga	120
ggtgcgggag	ctgctggagg	cgggggcgct	gccaacgcga	ccgaatagtt	acggctggag	180
gccgatccag	gtcatgatga	tgggcagcgc	ccgagtggcg	gagctgctgc	tgctccacgg	240
cgcggagccc	aactgcgcgg	accccgccac	tctcaccgga	cccgtgcacg	acgctgcccg	300
ggagggtctt	ctggacacgc	tgggtggtgct	gcaccggggc	ggggcgcgcc	tggacgtgcg	360
cgatgcctgg	ggcgtctgc	ccgtggacct	ggctgaggag	ctggggccatc	gcgatgtcgc	420
acgttacctg	cgcgcggtcg	cggggggcac	cagaggcagt	aaccatgccc	gcatagatgc	480
cgcggaaggt	ccctcagaca	tccccgattg	aaagaaccag	agaggctctg	agaaacctcg	540
ggaaacttag	atcatcagtc	accgaaggtc	ctacaggggc	acaactgccc	ccgccacaac	600

```
ccaccccgct ttcgtagttt tcatttagaa aatagagett ttaaaaatgt cctgcctttt 660
aacgtagata taagccttcc ccactaccg taaatgtcca tttatatcat tttttatata 720
ttcttataaa aatgtaaaaa agaaaaacac cgcttctgcc ttttactgt gttggagttt 780
tctggagtga gcactcacgc cctaagcgca cattcatgtg ggcatttctt gcgagcctcg 840
cagcctccgg aagctgtcga ctcatgaca agcattttgt gaactaggga agctcagggg 900
ggttactggc ttctcttgag tcacactgct agcaaattggc agaaccaaaag ctcaaataaa 960
aataaaataa ttttcattca ttcactc 987
```